TRANSCRIPTIONAL REGULATION OF THE HEPATITIS B VIRUS LARGE SURFACE ANTIGEN GENE

Anneke K. Raney

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, University of Warwick

Department of Molecular and Experimental Medicine,
The Scripps Research Institute

July, 1993

TABLE OF CONTENTS

TABLES AND FIGURES

TABLES

FIGURES

ACKNOWLEDGEMENTS

I am indebted to Dr. Andrew J. Easton for providing me with the opportunity to undertake this endeavor, and I am grateful to him for his valuable advice, his constant support, and his help with every aspect of completing this project.

I am very grateful to Dr. Alan McLachlan for his guidance and encouragement, his constructive criticism and advice, and his time and patience. I am also thankful to Dr. McLachlan for the provision of laboratory and technical facilities throughout the course of this project.

DECLARATION

I hereby declare that this thesis has been composed by myself, that it has not been accepted in any previous application for a degree, and that the work of which it is a record has been performed by myself. All sources of information have been specifically acknowledged by means of references. For ease of reference, the materials and methods section contains details of plasmid constructions and experimental techniques performed by others; where this occurs it is acknowledged in the text.

Anneke K. Raney

III. ABBREVIATIONS

In addition to the accepted Standard Abbreviations, the following were used.

A	Absorbance
ATP	Adenosine 5'-triphosphate
bp	Base pairs
C promoter	HBV nucleocapsid promoter
CAT	Chloramphenicol acetyltransferase
cpm	Counts per minute
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxyribonucleic acid triphosphates
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraaceticacid
Eh	HBV enhancer I
g	Gravity
GTF	General transcription factor
HBcAg	Hepatitis B virus core antigen
HBeAg	Hepatitis B virus e antigen
HBsAg	Hepatitis B virus surface antigen
HBV	Hepatitis B virus
HBx	Hepatitis B virus X protein
HNF1	Hepatocyte nuclear factor 1
kb	Kilobase
kbp	Kilobase pairs
kD	Kilodaltons

LUC	Luciferase
mRNA	Messenger RNA
nm	Nanometer
nt	Nucleotide(s)
ORF	Open reading frame
pA	Polyadenylation recognition sequence
PBS	Phosphate buffered saline
PS(1) promoter	HBV large surface antigen [preS(1)] promoter
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolutions per minute
S promoter	HBV major surface antigen promoter
SDS	Sodium dodecyl sulfate
SV40	Simian virus 40
TAF	TATA binding protein-associated factor
TBP	TATA binding protein
TLC	Thin layer chromatography
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer ribonucleic acid
vHNF1	Variant hepatocyte nuclear factor 1
vol	Volume
wt	Weight
X promoter	HBV X gene promoter

IV. SUMMARY

Hepatitis B virus (HBV) is a hepatotropic virus of highly restricted host range and tissue specificity. Although the mechanisms governing this tropism are not fully understood, it is likely that restrictions occur at multiple steps in the viral life cycle. The liver-specific regulation of HBV gene expression suggests that transcription may be an important factor in the hepatotropism of the virus. An analysis of tissue- or cell-line-specific regulation of the HBV promoters may elucidate the role of transcriptional regulation in the hepatotropism of the virus.

The major aim of this project was to characterize the transcriptional regulation of the large surface antigen gene of hepatitis B virus. To achieve this, the regions of the HBV genome involved in the regulation of the expression of the large surface antigen gene were identified using a transient transfection system in mammalian cell lines. The transcriptional activities of the four HBV promoters were compared in the human differentiated hepatoma cell lines Hep3B, PLC/PRF/5, HepG2 and Huh7, a human dedifferentiated hepatoma cell line HepG2.1, and the nonhepatoma cell lines HeLa S3 and NIH 3T3. To determine the relative transcriptional activities of the four HBV promoters, reporter gene plasmids were generated such that the expression of the firefly luciferase gene was under the control of each of the HBV promoters in the context of the complete genome. The nucleocapsid promoter and large surface antigen promoter displayed higher relative activities in the differentiated hepatoma cell lines, indicating that these promoters are preferentially active in these cell lines. A series of large surface antigen promoter deletion plasmids were constructed to identify the important regulatory regions of the large surface antigen promoter. The deletion analysis demonstrated that the region responsible for the high relative activity in differentiated hepatoma cell lines is located between -90 and -76 relative to the transcription initiation site (+1) located at map position 2809. This sequence element contains the binding site (GTTAATCATTACT) for the liver-enriched transcription factor hepatocyte nuclear factor 1, HNF1. A eukaryotic expression vector containing the HNF1 cDNA under the control of the mouse metallothionein I promoter was cotransfected with the HBV promoter constructs in Huh7 and HepG2.1 cells, and the relative levels of activity were determined. The Huh7 cell line was used because it is one of the cell lines in which HBV replication and particle production can occur and may represent the tissue culture system closest to the natural environment for the HBV life cycle, the liver cell. The cloned transcription factor HNF1 activated transcription from the large surface antigen promoter, but not from any of the other HBV promoters. Cotransfection experiments using the HNF1 cDNA expression vector and large surface antigen promoter deletion constructs demonstrated that this transactivation was mediated through the HNF1 binding site located between -90 and -76 in the large surface antigen promoter. A series of deletion mutants of the cDNA in the HNF1 expression vector was

generated to determine the transcriptional activation domain of the HNF1 polypeptide. The major domain of the HNF1 polypeptide involved in transcriptional activation of the large surface antigen promoter in the human hepatoma cell line HepG2.1 was mapped to a region rich in glutamine and proline residues (9 of 18 residues). To demonstrate directly that the HNF1 polypeptide produced by the expression of the HNF1 cDNA could bind the large surface antigen promoter HNF1 recognition sequence, and to determine whether a protein present in the differentiated hepatoma cell line Huh7 bound the HNF1 element, gel mobility shift analysis was performed. This analysis demonstrated that a protein present in nuclear extracts from Huh7 cells formed a specific complex with the HNF1 binding site which had similar migration properties to the complex formed between exogenously expressed HNF1 and the HNF1 recognition sequence. DNase I footprinting analysis demonstrated the binding of a protein present in the differentiated hepatoma cell line Huh7 to the HNF1 recognition sequence in the large surface antigen promoter. DNase I footprinting also showed that purified TATA binding protein binds the TATA box element located between -31 and -25 in the large surface antigen promoter. The analysis of synthetic promoter constructs suggested that the HNF1 and TATA box elements were the only elements necessary for maximal activity from the large surface antigen promoter, and analysis of clustered point mutations in the large surface antigen minimal promoter region demonstrated that sequences between the HNF1 and TATA box elements were not required for the HNF1-dependent activity of the large surface antigen promoter. These studies suggested that the liver-enriched transcription factor HNF1 plays a critical role in the cell-line and tissue-specific regulation of the HBV large surface antigen promoter.

V. INTRODUCTION

V. A. The Biology of Hepatitis B Virus

Introduction

Hepatitis B virus is the prototype member of a family of animal viruses related in morphology, genetic organization, life cycle and tissue tropism. These viruses are small, circular, partially double-stranded DNA viruses which are primarily hepatotropic. The unique features of this group determined its classification as a separate virus family, the *hepadnaviridae*, so named for its hepatotropism and DNA genome (Gust *et al.*, 1986). Other members of this family include the woodchuck hepatitis virus (WHV) (Summers *et al.*, 1978), Beechey ground squirrel hepatitis virus (GSHV) (Marion *et al.*, 1980), Pekin duck hepatitis B virus (DHBV) (Mason *et al.*, 1980), and heron hepatitis B virus (HHBV) (Sprengel *et al.*, 1988). Another possible member of this group is the tree squirrel hepatitis B virus (TSHV) (Feitelson *et al.*, 1986). However, the genome of TSHV has not yet been cloned, so its identity as a virus distinct from GSHV has not been definitively established.

History of HBV

Following the discovery of Australia (Au) antigen, or serum hepatitis (SH) antigen as it was also called, in human serum by Blumberg and colleagues in 1965 (Blumberg *et al.*, 1965; Prince, 1968b), the infectious agent responsible for viral hepatitis was identified and characterized (Robinson & Greenman, 1974; Robinson, 1977; Robinson *et al.*, 1974; Kaplan *et al.*, 1973; Hruska *et al.*, 1977; Landers *et al.*, 1977). Studies revealed that Au antigen was frequently found in sera of hepatitis patients and that Au antigen-positive sera contained particles that reacted with antibodies to the Au antigen (Bayer *et al.*, 1968; Sutnick *et al.*, 1967; Sutnick *et al.*, 1968; Prince, 1968a; Blumberg *et al.*, 1969). The Au antigen, now known as the envelope or hepatitis B surface antigen (HBsAg), was present on spheres and filaments with 22 nanometer (nm) diameter and length varying from less than 50 nm to 1000 nm (Bayer *et al.*, 1968; Dane *et al.*, 1970; Almeida & Waterson, 1969; Bond & Hall, 1972). Subsequently a 42 nm

sphere, the Dane particle, was observed in the sera of serum hepatitis patients, and based on its morphology was suggested to be the agent responsible for the disease (Robinson, 1977; Dane *et al.*, 1970; Gust *et al.*, 1970; Jokelainen *et al.*, 1970).

Hepatitis B Virus Particles

The identification of the Dane particle as the infectious virion resulted from the work of Robinson and colleagues (Robinson & Greenman, 1974; Robinson, 1977; Robinson *et al.*, 1974; Kaplan *et al.*, 1973; Hruska *et al.*, 1977; Landers *et al.*, 1977). Disruption of the 42 um Dane particle with nonionic detergent removes the envelope containing HBsAg and cell-derived lipid and releases a 28 nm icosahedral nucleocapsid or core particle whose main component is a 21 kilodalton (kD) protein, the hepatitis B core antigen (HBcAg) (Almeida *et al.*, 1971) (Fig. V. 1). The nucleocapsid carries the viral genome, a 3.2 kilobase (kb) partially double-stranded DNA molecule, and a protein which is attached to the genome. In addition to the infectious Dane particles, 22 nm HBsAg spheres and filaments are secreted into the serum and are present in large excess compared with the Dane particles during HBV infection (Dane *et al.*, 1970) (Fig. V. 1). These subviral particles do not contain the viral DNA and are not infectious. The HBV Dane particles, filaments and spheres are composed of the same HBsAg polypeptides but they contain different ratios of each. The Dane particles and filaments contain approximately 10 to 20% of each of the middle and large polypeptides, with the remaining 60 to 80% of the protein content contributed by the major surface antigen polypeptide (Heermann *et al.*, 1984). The 22 nm spheres contain only 1 to 2% large, 10 to 20% middle, and 78 to 89% major surface antigen polypeptide, with a total of about 100 polypeptide subunits (Heermann *et al.*, 1984). It is interesting to note that expression studies of the surface antigens have demonstrated that the middle and major surface polypeptides can assemble into 22 nm spheres (Moriarty *et al.*, 1981; Stenlund *et al.*, 1983; Hsiung *et al.*, 1984; Michel *et al.*, 1984; Persing *et al.*, 1985; McLachlan *et al.*, 1987; Ou & Rutter, 1987; Cheng & Moss, 1987; Molnar-Kimber *et al.*, 1988; Langley *et al.*, 1988; Yoneyama *et al.*, 1988), but assembly of 22 nm filaments requires the coproduction of large surface antigen polypeptides (McLachlan *et al.*, 1987; Chisari *et*

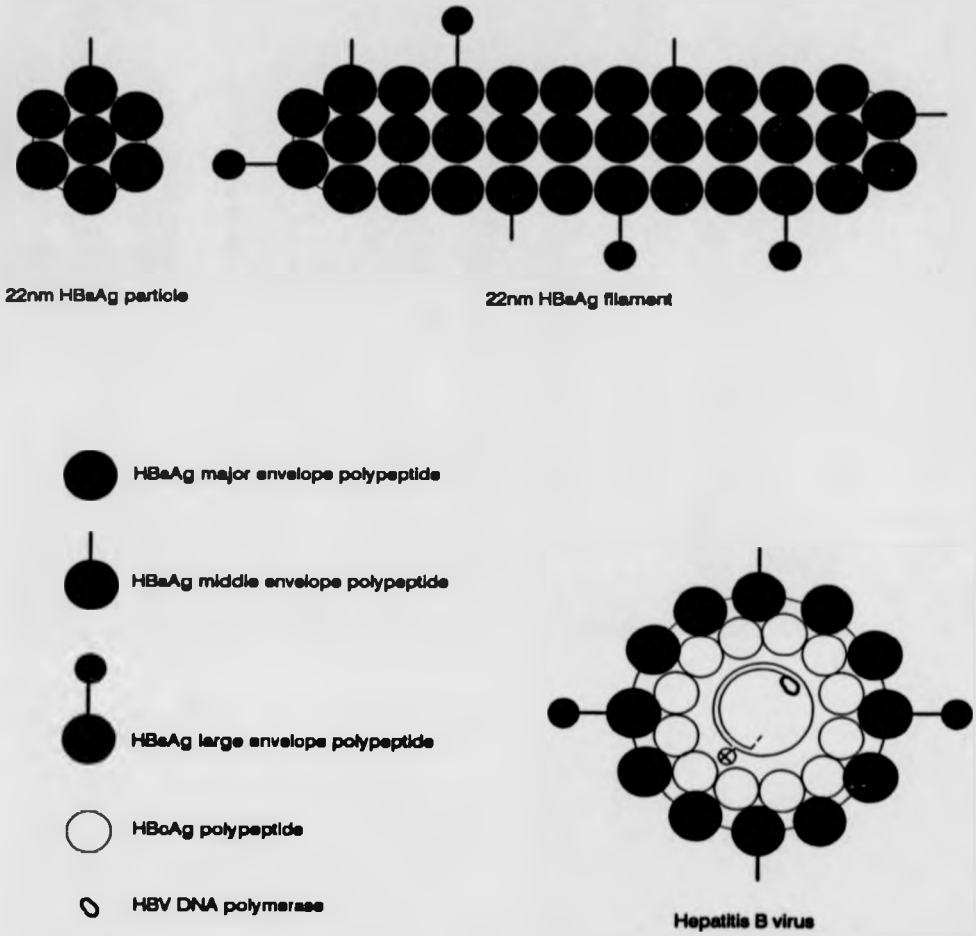


Figure V.1. Hepatitis B virus (HBV) particles. Diagrammatic representation of the structure of the 22 nm hepatitis B surface antigen (HBsAg) sphere, 22 nm filament, and hepatitis B virus (Dane particle). HBcAg; hepatitis B core antigen.

al., 1987), indicating that the large envelope polypeptide influences assembly of the viral particles. It has been shown that the large surface antigen is required for the assembly of the virion, demonstrating its critical importance in the life cycle of the virus (Ueda *et al.*, 1991).

HBV Infection

HBV infection is a worldwide health problem and is endemic in many regions of Asia and Africa (Nishioka *et al.*, 1975). Maintenance of HBV infection in the population occurs by both vertical and horizontal transmission of the virus (Gust *et al.*, 1986). Vertical transmission of infection occurs by the passage of HBV from an infected mother to her child at or prior to birth by mechanism(s) which are currently unclear (Lee *et al.*, 1978). Horizontal transmission of HBV infection can occur during sexual contact between infected and uninfected individuals and by percutaneous contact (e.g. injection) or other close contacts between the virus and man (e.g. mother to infant and between patient and healthcare worker) (Gust *et al.*, 1986; Prince *et al.*, 1970; Lewis *et al.*, 1973; Szmuness *et al.*, 1975; Dietzman *et al.*, 1977; Lauer *et al.*, 1979; Tong *et al.*, 1981).

A consequence of infection with HBV is the induction of the host immune response by the various viral antigens. These antigen-antibody responses include the induction of anti-HBs by the major surface antigen polypeptide (Ganem, 1982), anti-preS2 by the preS2 region of the middle and large surface antigen polypeptides (Neurath *et al.*, 1985; Neurath *et al.*, 1986a; Alberti *et al.*, 1988; Brahm *et al.*, 1988), anti-preS1 by the preS1 region of the large surface antigen polypeptide (Neurath *et al.*, 1985; Klinkert *et al.*, 1986), anti-HBc by the HBcAg (Ganem, 1982), anti-HBe by the HBeAg (Ganem, 1982), anti-HBx by the X gene polypeptide (Katayama *et al.*, 1989; Moriarty *et al.*, 1985; Kay *et al.*, 1985; Meyers *et al.*, 1986; Elfassi *et al.*, 1986; Liang *et al.*, 1988; Katayama *et al.*, 1989; Levrero *et al.*, 1990), and anti-HBp by the DNA polymerase polypeptide (Stemler *et al.*, 1988; Feitelson *et al.*, 1988; Weimer *et al.*, 1989; Chang *et al.*, 1989a).

The first evidence of primary infection with HBV during a self limited HBsAg positive infection is the observation of HBsAg in serum (Ganem, 1982). This is

followed by the appearance of HBV virus particles with concomitant endogenous DNA polymerase activity and HBeAg antigenemia. As the levels of these viral particles and polypeptides in the serum start to decline, anti-HBc appears in the serum. At this time clinical hepatitis can be observed as elevated levels of liver transaminases in the serum. The decline of the clinical symptoms is followed by a rise in anti-HBs and anti-HBe titers in the serum which subsequently decrease with time. The observation of these antiviral antibody responses indicates the resolution of the infection. Self limiting HBV infection without detectable HBsAg can occur and is indicated by observable anti-HBs and anti-HBc titers in the serum in the absence of detectable HBV antigenemia (Ganem, 1982).

The course of a persistent or chronic HBV infection is initially similar to the HBsAg positive self-limited infection (Ganem, 1982). HBsAg, HBeAg, viral particles and anti-HBc appear in the serum. The infection may be asymptomatic or characterized by liver damage. The infection fails to resolve, and circulating HBsAg and virus particles generally persist for years without seroconversion to an anti-HBs status. After an extended period of chronic infection, viral HBV DNA can often be found integrated into the host chromosomal DNA (Shafritz *et al.*, 1981; Takada *et al.*, 1990). During this period of time seroconversion from HBeAg-positive to anti-HBe-positive may occur. The consequence of this event can be the generation of a chronic carrier state where HBsAg antigenemia exists in the absence of serum HBV particles (Realdi *et al.*, 1980; Hoofnagle *et al.*, 1981).

HBV infection can lead to hepatocellular necrosis in man (Dienstag *et al.*, 1982) though the mechanism of liver damage and the factors determining the severity and outcome of infection are not fully understood. However, the observation that hepatocytes of asymptomatic chronic carriers can produce large quantities of HBV without apparent liver disease indicates that the virus is not directly cytopathic. In support of this conclusion, cell lines and transgenic mice producing HBV particles have demonstrated no cytopathic effect due to virus production (Sureau *et al.*, 1986; Tsurimoto *et al.*, 1987; Sells *et al.*, 1987; Chang *et al.*, 1987; Yaginuma *et al.*, 1987; Araki

et al., 1989; Shih *et al.*, 1989; Farza *et al.*, 1988). In addition, several observations implicate the host cellular immune response as a major factor contributing to the pathogenesis of liver disease. These include (i) infected liver tissue reveals mononuclear inflammatory cells in close proximity to necrotic hepatocytes, (ii) immunomodulation of infected patients affects the severity of liver damage, (iii) the serology during and after infection correlates with the observed temporal elevation in serum transaminase levels, and (iv) agammaglobulinemic patients can show liver damage during HBV infection (Dudley *et al.*, 1972; Alberti *et al.*, 1984; Dienstag, 1984).

HBsAg chronic carrier status has additional clinical implications. The estimated relative risk of primary hepatocellular carcinoma (PHC) in these individuals is approximately 100 times greater than in those serologically negative for HBsAg (Beasley *et al.*, 1981; Beasley, 1988). In addition, tumors and cell lines derived from chronic HBV carries usually contain HBV DNA integrated in the chromosomal DNA (Ganem & Varmus, 1987), suggesting a possible role for the virus in the transformation process occurring in the infected hepatocyte. In one tumor, HBV sequences have been observed integrated into a homologue of the known cellular proto-oncgene, *c-erbA*, suggesting a direct role for viral DNA in the transformation process (Dejean *et al.*, 1986). However, in the majority of tumors which have been examined, the integration sites for HBV DNA are at different chromosomal locations suggesting that the role of these sequences in the transformation process is likely to be more complex than the simple insertional mutagenesis models (Ganem & Varmus, 1987).

A putative variant form of HBV, HBV-2, has been described (Coursaget *et al.*, 1987). Infection with this virus is characterized by the presence of serum HBsAg but the absence of anti-HBc. The subsequent loss of HBsAg occurs without the appearance of HBeAg or the production of anti-HBs or anti-HBc (Coursaget *et al.*, 1987). Similar serological evidence has been reported from several regions of the world, suggesting this infectious agent may be relatively widespread (Budkowska *et al.*, 1988; Redeker & Govindarajan, 1988; Wu *et al.*, 1988; Tagariello *et al.*, 1989; Echevarria *et al.*, 1988; Wands *et al.*, 1986). HBV-2 has not been characterized at the molecular level, and the precise relationship to HBV is currently unknown.

Structure of the Virus Particle

The envelope of the infectious Dane particle consists of viral surface antigen polypeptides anchored in a cell-derived lipid bilayer. The envelope contains approximately 400 HBsAg subunits, of which 40 to 80 are middle surface antigen polypeptide, 40 to 80 are large surface antigen polypeptide, and the remainder are major surface antigen polypeptides. The envelope surrounds a 28 nm nucleocapsid containing the viral genome. The nucleocapsid contains approximately 180 hepatitis B core antigen (HBcAg) polypeptides which assemble into the icosahedral nucleocapsid structure (Pasek *et al.*, 1979; Heermann *et al.*, 1984; Budkowska *et al.*, 1977; Hruska & Robinson, 1977; Petit & Pillot, 1985; Theilmann *et al.*, 1989a; Theilmann *et al.*, 1989b; Yamaki *et al.*, 1982; Onodera *et al.*, 1982). HBcAg is a 21 kD phosphoprotein which may possess autophosphorylating serine protein kinase activity (Gerlich *et al.*, 1982; Albin & Robinson, 1980; Roossinck & Siddiqui, 1987). The nucleocapsid displays endogenous DNA polymerase activity, which is probably encoded by the viral polymerase open reading frame (ORF) (Toh *et al.*, 1983; Bavand *et al.*, 1989; Mack *et al.*, 1988; Kaplan *et al.*, 1973; Robinson & Greenman, 1974). The nucleocapsid also contains the 3.2 kilobase (kb), partially double-stranded circular DNA genome. This genome is one of the smallest of any known animal DNA virus and contains several interesting features. The genome is not completely double stranded. The full-length, or long (L) strand is complementary to the viral mRNA and is therefore also called the minus (-) strand. The opposite strand, the plus (+) or short (S) strand, is less than unit length due to the variability of the 3' terminus. The position of the 3' terminus varies depending upon the extent of synthesis of the plus strand during the viral life cycle as described below. The genome of HBV therefore contains a single-stranded region, from 15% to 50% of its length (Hruska *et al.*, 1977; Landers *et al.*, 1977; Summers *et al.*, 1975). The circularity of the genome is maintained by a 5'-terminal cohesive overlap of the two DNA strands approximately 226 nucleotides in length (Sattler & Robinson, 1979). The long strand also contains a nick or short gap and a protein, which is probably encoded by the ORF which also encodes the viral polymerase ORF, covalently attached at the 5' terminus

(Sattler & Robinson, 1979; Gerlich & Robinson, 1980). The short strand is believed to contain an oligoribonucleotide derived from the viral pregenomic RNA attached at the 5' terminus, as observed in DHBV, GSHV, and WHV (Lien *et al.*, 1986; Will *et al.*, 1987; Seeger *et al.*, 1986). The unique characteristics of the 5' termini reflect their roles in the replication of the virus.

Genome Organization

The cloning and sequencing of the HBV genome from purified Dane particles have revealed the genetic organization of HBV (Valenzuela *et al.*, 1979; Galibert *et al.*, 1979; Pasek *et al.*, 1979; Ono *et al.*, 1983) (Fig. V. 2). This organization is similar in the related animal viruses. Examination of the sequence reveals a compact coding capacity of the virus. Each nucleotide is contained in an ORF, with 50% of the nucleotides present in more than one potential coding region. Several conserved *cis*-acting regulatory sequence elements have also been observed in the sequences of the various subtypes. The conserved *cis*-acting sequences present in the PS(1) promoter are indicated in Fig. V. 3. Sequence comparison of the major subtypes reveals some highly conserved regions important in the life cycle of the virus (Raney & McLachlan, 1991). Four major ORFs conserved between all of the viral genomes are transcribed from the full-length minus strand DNA. These ORFs encode the envelope or surface antigens (HBsAg) (Peterson, 1981; Peterson *et al.*, 1977; Stibbe & Gerlich, 1983; Heermann *et al.*, 1984), the nucleocapsid antigens (HBcAg and HBeAg) (Pasek *et al.*, 1979; Takahashi *et al.*, 1983; Standring *et al.*, 1988), the DNA polymerase gene product (Toh *et al.*, 1983; Bavand *et al.*, 1989; Mack *et al.*, 1988), and the X gene product (Chang *et al.*, 1987; Chisaka *et al.*, 1987; Katayama *et al.*, 1989). Additional ORFs have been identified by computer analysis in HBV and WHV (Kaneko & Miller, 1988; Chen *et al.*, 1993) but there is no evidence to suggest that these represent authentic genes.

The surface antigen ORF is divided into the S region, the preS(2) region and the preS(1) region by three in-frame translation initiation codons at map positions 157, 3174, and 2850, respectively (subtype *ayw*). The translation products of this gene are the major (or small), middle and large HBsAg polypeptides of 25, 31, and 43 kD,

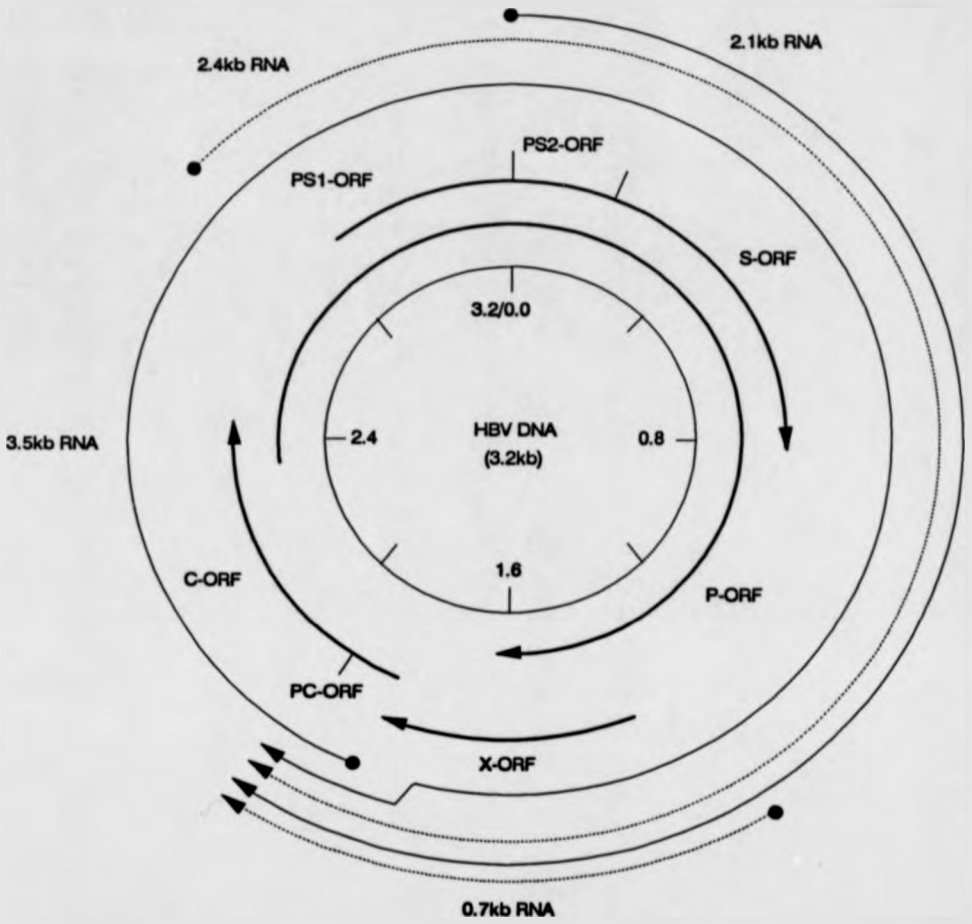


Figure V.2. Organization of the HBV genome. The coordinates of the HBV DNA are numbered using the first A residue of the *Eco*RI site as nucleotide 1. The open reading frames are designated by thick-lined arrows. The RNA transcripts are designated by dotted or thin-lined arrows. PS, presurface; S, surface; P, polymerase; X, X gene; PC, precore; C, core; ORF, open reading frame.

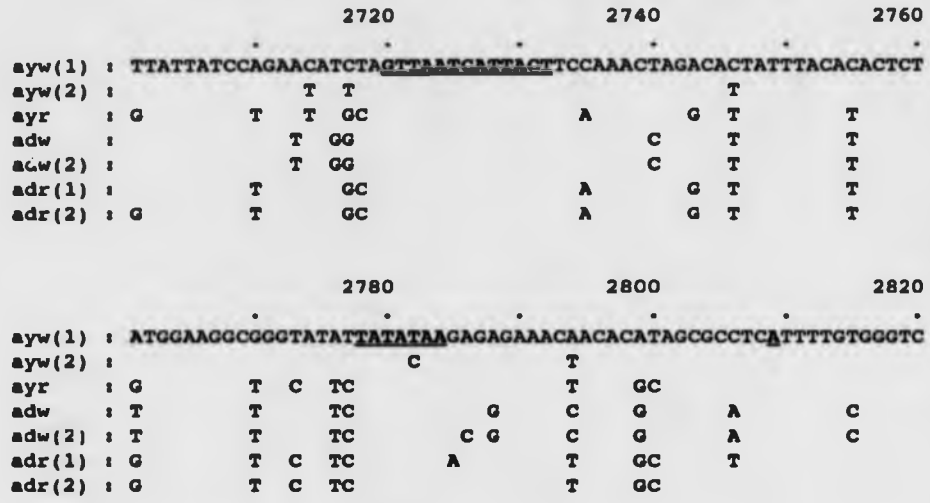


Figure V.3. Comparison of the HBV DNA sequences of the major subtypes in the PS(1) promoter region. The HBV genomes are subtypes *ayw* [sequences 1 (Galibert *et al.*, 1979) and 2 (Bichko *et al.*, 1985)], *ayr* (Okamoto *et al.*, 1986), *adw* (Ono *et al.*, 1983), *adw*$_2$ (Valenzuela *et al.*, 1980), and *adr* [sequences 1 (Kobayashi & Koike, 1984) and 2 (Ono *et al.*, 1983)]. The sequences were aligned with the *ayw*$_1$ sequence using the first A residue of the *Eco*RI site as nucleotide 1. The differences between the *ayw*$_1$ sequence and the other sequences are indicated. The hepatocyte nuclear factor 1 (HNF1) and TATA binding sequences are indicated by bold, underlined letters. The transcription initiation site for the large surface antigen mRNA located at nt 2809 is indicated by a bold, underlined letter.

respectively (Galibert *et al.*, 1979) (Fig. V. 4). These polypeptides share a common carboxyl terminus and differ by the addition of amino-terminal regions depending on the translation initiation codon used for their synthesis. The 25 kD major surface antigen (HBsAg/P25) comprises 226 amino acids and can be glycosylated to generate HBsAg/GP28 (Heermann *et al.*, 1984). The middle surface antigen contains an additional 55 amino acids at the amino terminus of the major surface antigen and is variably glycosylated, giving rise to the polypeptides HBsAg/GP33 and GP36 (Heermann *et al.*, 1984). Translation from the first in-frame initiation codon results in the synthesis of the large surface antigen (HBsAg/P43), which also exists in a glycosylated form (HBsAg/GP46) (Heermann *et al.*, 1984). The large surface antigen polypeptide differs from the middle surface antigen by the addition of 108 or 119 amino-terminal amino acids encoded by the preS(1) region. The exact length of the addition depends on the viral strain.

The nucleocapsid ORF contains two in-frame translation initiation codons at map positions 1816 and 1903. The 29-amino acid region between the two initiation codons is called the precore region (Fig. V. 5). Translation from the downstream AUG produces the core (HBcAg) polypeptide, a 21 kD protein (Pasek *et al.*, 1979). Translation from the upstream, precore AUG potentially yields a 24 kD protein, but the polypeptide is proteolytically processed at both the amino and carboxyl termini to produce an 18 kD secreted protein, HBeAg (Takahashi *et al.*, 1983; Standring *et al.*, 1988). The largest of the four major ORFs, spanning approximately 75% of the genome, has the capacity to encode a polypeptide of 94 kD. This ORF encodes the multifunctional viral polymerase protein, which includes a terminal protein attached to the minus strand of the genome in the virion, an RNase H activity, reverse transcriptase activity and DNA polymerase activity (Radziwill *et al.*, 1990; Bartenschlager & Schaller, 1988). The X gene is the smallest of the four ORFs and potentially encodes a polypeptide of 17 kD.

The genome organization of the woodchuck and ground squirrel hepatitis viruses is similar to that of HBV, with four overlapping open reading frames. These ORFs encode the structural surface and nucleocapsid proteins, the viral polymerase and the X

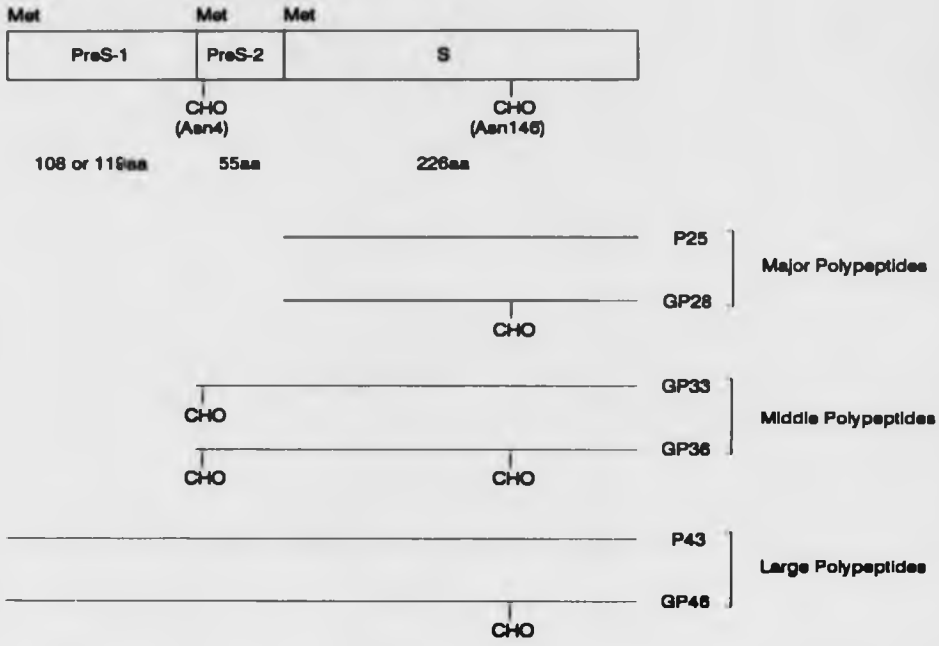


Figure V.4. Organization of the hepatitis B virus surface antigen (HBsAg) open reading frame. The methionine (Met) residues encoded by the initiation codons for the large, middle, and major surface antigen polypeptides are indicated. The asparagine (Asn) residues at which carbohydrate (CHO) can modify the HBsAg are shown. PreS, presurface; S, surface.

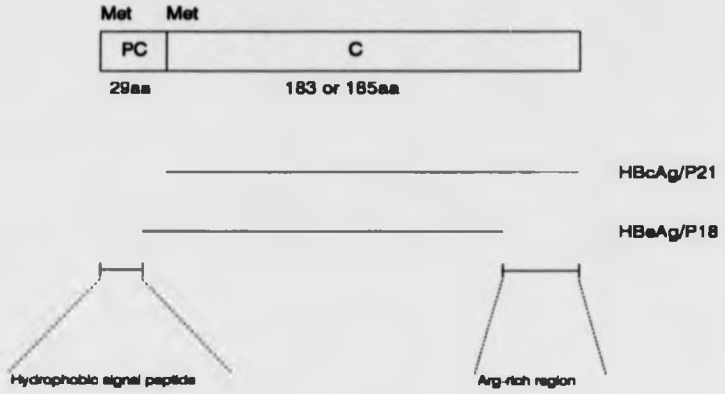


Figure V.5. Organization of the nucleocapsid open reading frame. The methionine (Met) residues encoded by the precore (PC) and core (C) initiation codons are indicated. The 19-amino acid hydrophobic signal peptide and the 34-amino acid arginine-rich carboxy-terminal region cleaved from the HBeAg/P24 precursor to produce the secreted HBeAg/P18 are shown.

gene product. They lie in the same relative position as in HBV and employ multiple translation initiation codons to produce multiple polypeptides from single ORFs. The two major RNA molecules produced by these viruses are analogous to those of HBV. Both viruses have a "pregenomic" RNA (3.7 kb for WHV, 3.5 kb for GSHV) and a subgenomic 2.1 kb (WHV) or 2.3 kb (GSHV) RNA. Translation from these mRNAs could produce the nucleocapsid and polymerase polypeptides and the surface antigen polypeptides. An X gene mRNA has been identified in WHV but not in GSHV.

The organization of the duck hepatitis genome is similar to that of the mammalian hepadnaviruses except that three, rather than four, ORFs represent the coding regions of the viral DNA (Sprengel *et al.*, 1985). The fourth ORF, designated X in the mammalian hepadnaviruses, is absent from the DHBV genome.

Viral RNAs

The major HBV RNAs are unspliced, polyadenylated and 3' coterminal (Fig. V. 2). These RNA molecules have been studied in a number of systems including human cell lines containing integrated viral sequences (Ou & Rutter, 1985; Edman *et al.*, 1980; Chakraborty *et al.*, 1980), transfection of cloned HBV DNA into rodent and human cell lines (Pourcel *et al.*, 1982; Cattaneo *et al.*, 1983; Standring *et al.*, 1984; Siddiqui *et al.*, 1986; Sureau *et al.*, 1986; Tsurimoto *et al.*, 1987; Zelent *et al.*, 1987b; Sells *et al.*, 1988), infected chimpanzee liver (Cattaneo *et al.*, 1984), transgenic mice (Babinet *et al.*, 1985; Farza *et al.*, 1988; Araki *et al.*, 1989), and *in vitro* transcription systems (Chakraborty *et al.*, 1981; Rall *et al.*, 1983). The two predominant RNA species present during HBV infection are the pregenomic 3.5 kb and subgenomic 2.1 kb molecules (Cattaneo *et al.*, 1983; Cattaneo *et al.*, 1984; Yokosuka *et al.*, 1986; Imazeki *et al.*, 1987; Su *et al.*, 1989b). The 3.5 kb RNA is called "pregenomic" because of its role as the template for genomic DNA synthesis as well as its function as a messenger RNA. The 5' ends of the 3.5 kb RNAs are heterogeneous. Multiple transcription initiation sites span the initiation codon for the precore signal sequence which precedes the HBcAg ORF (Yaginuma *et al.*, 1987; Sells *et al.*, 1988). RNA transcripts initiated upstream of the precore AUG encode the nonparticulate, secreted 18 kD e antigen (HBeAg) and those initiated downstream of the

precore initiation codon encode the 21 kD core polypeptide. The pregenomic RNA also probably encodes the polymerase gene product as it is the only identified transcript large enough to do so.

The 2.1 kb RNA also has multiple transcription initiation sites which span the translation initiation codon for the middle envelope polypeptide (the preS(2) region) (Yaginuma *et al.*, 1987; Sells *et al.*, 1988; Standring *et al.*, 1984) and terminates at the single polyadenylation site in the HBV genome (Sells *et al.*, 1988; Simonsen & Levinson, 1983; Standring *et al.*, 1984). These mRNAs encode the major and middle surface antigen polypeptides. The expression of the only other known HBV structural polypeptide, the large envelope polypeptide, occurs from a 2.4 kb mRNA which initiates approximately 40 nucleotides (nt) upstream of the translation initiation site for this ORF (Yaginuma *et al.*, 1987; Sells *et al.*, 1988; Farza *et al.*, 1988). Although the evidence for the existence of this mRNA in infected liver tissue is limited (Will *et al.*, 1987; Cattaneo *et al.*, 1984), it has clearly been identified in human hepatoma cell lines which are synthesizing HBV particles (Sureau *et al.*, 1986; Tsurimoto *et al.*, 1987; Chang *et al.*, 1987; Sells *et al.*, 1988).

In addition, it is likely that a 0.7 kb mRNA initiated upstream of the X gene open reading frame is synthesized at some stage during HBV infection. Although there is no direct evidence for the synthesis of this RNA molecule in human infections, a 0.7 kb RNA derived from the X gene region of the HBV genome has been observed in transgenic mice replicating HBV (Araki *et al.*, 1989) and in cell lines transfected with HBV DNA (Gough, 1983; Simonsen & Levinson, 1983; Siddiqui *et al.*, 1986; Saito *et al.*, 1986; Siddiqui *et al.*, 1987). In addition, an equivalent RNA species has been identified in RNA isolated from woodchucks infected with WHV (Kaneko & Miller, 1988).

Life Cycle of the Virus

The inability to infect permanent tissue culture cell lines with HBV has been a major obstacle to identifying the mechanism of viral entry into hepatocytes (Hirschman, 1984; Rijntjes *et al.*, 1988; Gripon *et al.*, 1988) and has prevented the identification of putative viral receptors and their ligands. However, evidence has been presented

suggesting there may be selective interactions between surface antigen polypeptides and cell surface elements of several hepatic and non-hepatic tissue culture cell lines and also between the envelope polypeptides and primary human hepatocyte plasma membrane components (Neurath *et al.*, 1986b; Peeples *et al.*, 1987; Komai *et al.*, 1988; Pontisso *et al.*, 1989a; Pontisso *et al.*, 1989b). The exact nature of these interactions is unknown, but they may be direct, receptor-ligand interactions or more indirect interactions through polyalbumin, the albumin receptors on the surface of the human hepatocytes (Weisiger *et al.*, 1981; Ockner *et al.*, 1983; Thung & Gerber, 1981), and the receptor for polymerized human serum albumin present in the preS(2) region of the envelope polypeptides (Michel *et al.*, 1984; Persing *et al.*, 1985; Machida *et al.*, 1984). The role of these interactions in the infection of hepatocytes and possibly other cell types is currently unclear.

After entry into the hepatocyte, presumably by receptor mediated endocytosis or fusion with the plasma membrane, the virus releases the nucleocapsid from the surface antigen envelope. (A diagram summarizing the viral life cycle is shown in Fig. V. 7). Whether this process is linked with virus entry or represents a separate step in the viral life cycle is not clear. The viral genome is then translocated to the nucleus, presumably within the nucleocapsid and mediated by the nuclear translocation signal present in the HBcAg polypeptide (McLachlan *et al.*, 1987). Several modifications of the infecting viral genome must be made before transcription of the viral mRNAs can occur (Will *et al.*, 1987). The single-stranded region of the HBV genome is converted to double-stranded DNA probably by the endogenous viral DNA polymerase. It is known that the synthesis of this region of DNA can be completed *in vitro* by the endogenous HBV DNA polymerase present in the nucleocapsid of the virus (Robinson & Greenman, 1974; Lutwick & Robinson, 1977; Hruska *et al.*, 1977). Therefore, it is likely that during infection the same enzyme converts the partially double-stranded DNA into nicked circular double-stranded DNA. This event could occur within the nucleocapsid or in the nucleoplasm after the genome has been released from the nucleocapsid. In addition to the synthesis of DNA, other modifications necessary to generate the transcriptional

template are the removal of the terminal protein and oligoribonucleotide attached to the 5' ends of the viral DNA strands, and the ligation of the nicks in the double-stranded DNA to generate a covalently closed, circular, double-stranded HBV genome (Miller & Robinson, 1984).

The covalently closed, circular, double-stranded HBV DNA molecule is probably the template for HBV transcription. HBV synthesizes two classes of RNA, genomic and subgenomic (Fig. V. 2). The 3.5 kb genomic RNAs contain all of the viral genetic information and therefore must serve as the replication template as well as messenger RNA. The subgenomic mRNAs include 2.1 kb and 2.4 kb, and possibly 0.7 kb messages. All of these viral mRNAs are unspliced, of plus-strand polarity, and polyadenylated at a common 3' terminus in the core region. In addition to the unspliced species, spliced HBV RNAs have been described (Simonsen & Levinson, 1983; Su *et al.*, 1989a; Chen *et al.*, 1989). Since it is not clear if these RNAs code for HBV polypeptides, their significance in the life cycle of HBV is currently unknown.

The transcription map of HBV clearly identifies the mechanisms for the production of the surface and core polypeptides. The three major mRNAs initiate in the regions immediately upstream of, and span the coding regions for, the surface and core polypeptides. Detection of these polypeptides has been correlated with the presence of these mRNAs. The way in which the products of the X coding region and the polymerase coding region are expressed is less clear. The surface antigen polypeptides contain signal sequences which direct their translocation into the endoplasmic reticulum membrane (Eble *et al.*, 1986; Eble *et al.*, 1987; Simon *et al.*, 1988). The major surface antigen promoter is considerably more active than the large surface antigen promoter. This is reflected by the relative abundance of the 2.1 kb and 2.4 kb RNA molecules and by the synthesis of a large excess of the major envelope polypeptide (Sureau *et al.*, 1986; Tsurimoto *et al.*, 1987; Chang *et al.*, 1987; Yaginuma *et al.*, 1987; Cattaneo *et al.*, 1984; Yokosuka *et al.*, 1986; Su *et al.*, 1989b; Raney *et al.*, 1990; Robinson & Lutwick, 1976). The middle surface antigen polypeptide is synthesized relatively inefficiently from the RNA transcripts expressed from the major surface antigen promoter for two reasons.

The translation initiation codon for the middle envelope polypeptide is located about 15 nucleotides from the 5' end of the longest 2.1 kb surface antigen RNA and is absent from the shorter 2.1 kb RNAs (Yaginuma *et al.*, 1987; Sells *et al.*, 1988; Standring *et al.*, 1984). In addition, the sequence around the translation initiation codon does not conform to the favored consensus sequence for efficient translational initiation (Stibbe & Gerlich, 1983; McLachlan *et al.*, 1987; Price *et al.*, 1988). This may explain why the middle envelope polypeptide is present at less than 20% the level of the major surface antigen polypeptide in the sera of infectious individuals (Stibbe & Gerlich, 1983; Heermann *et al.*, 1984; Stibbe & Gerlich, 1982). The envelope polypeptides appear to be expressed in excess over the other viral components. They have the capacity to assemble into 22 nm spheres and filaments which bud into the lumen of the endoplasmic reticulum and are subsequently secreted by passage through the endoplasmic reticulum and Golgi compartment (Eble *et al.*, 1986; Eble *et al.*, 1987; Simon *et al.*, 1988; Patzer *et al.*, 1984; Patzer *et al.*, 1986; Eble *et al.*, 1990). This process accounts for the large excess of subviral particles over virions in the sera of viremic individuals (Robinson & Lutwick, 1976). However, the reason subviral particles are synthesized in excess over virions is unknown.

The HBcAg/P21 polypeptide, which is the major component of the nucleocapsid, is produced by translation from the mRNAs initiated after the precore initiation codon. The HBeAg is produced from the translation product of the longest 3.5 kb viral RNA. It is generated from a precursor polypeptide containing a signal peptide encoded by the precore sequence which targets this polypeptide to the endoplasmic reticulum (Fig. V. 5) (Standring *et al.*, 1988; McLachlan *et al.*, 1987; Garcia *et al.*, 1988; Weimer *et al.*, 1987; Bruss & Gerlich, 1988; Jean-Jean *et al.*, 1989). After the cleavage of the amino-terminal signal sequence by the signal peptidase in the endoplasmic reticulum, the HBeAg/P22 polypeptide proceeds through the secretion pathway and the carboxy-terminal 34 amino acids are removed by proteolytic processing (Takahashi *et al.*, 1983), resulting in the secretion of the mature HBeAg/P18 polypeptide. The carboxyl-terminal processing probably occurs in the late or post-Golgi compartment by cellular proteases

possessing trypsin- and carboxypeptidase B-like specificities which could correctly modify the arginine-rich carboxy-terminal region of the unprocessed HBeAg/P22 polypeptide (Fig. V. 5) (Docherty & Steiner, 1982). As with the secretion of the large excess of the subviral envelope antigen particles, the function of circulating HBeAg in the serum of viremic patients is not known. However, it has been speculated that HBeAg might have a role in the modulation of the host immune response to the virus (Thomas *et al.*, 1988). In a transgenic mouse model, a role for HBeAg in neonatal tolerance has been demonstrated, supporting the hypothesis that maternal HBeAg from a chronic carrier mother can tolerize the fetus *in utero*, predisposing it to chronic infection upon exposure to HBV at birth (Milich *et al.*, 1990).

Although there is no direct evidence that the X gene is transcribed during HBV infection, this open reading frame is conserved between all of the mammalian hepadnaviruses, suggesting the functional importance of the X gene in the life cycle of the virus (Tiollais *et al.*, 1985). Antibodies to the X gene product have been observed in the sera of infected individuals indicating that the X gene polypeptide is synthesized during the course of at least some HBV infections (Siddiqui *et al.*, 1987; Moriarty *et al.*, 1985; Kay *et al.*, 1985; Meyers *et al.*, 1986; Elfassi *et al.*, 1986; Liang *et al.*, 1988; Katayama *et al.*, 1989; Levrero *et al.*, 1990). The role of the X gene polypeptide in the HBV life cycle is currently unclear. However, it appears that the product of the X gene can transactivate several heterologous promoters (Spandau & Lee, 1988; Twu & Schloemer, 1987; Zahm *et al.*, 1988; Seto *et al.*, 1988; Twu & Robinson, 1989; Siddiqui *et al.*, 1989; Colgrove *et al.*, 1989; Seto *et al.*, 1989) suggesting it may regulate the expression of various viral and cellular genes important in the production of viral particles.

Although the mechanism of expression of the HBV DNA polymerase gene has not been definitively characterized, it is known that the 3.5 kb RNAs are the only identified RNAs which can encode the polymerase polypeptide. It has been shown that the expression of the DHBV DNA polymerase is achieved by internal entry of ribosomes on the 3.5 kb RNA and initiation of translation from the first in-frame translation

initiation codon of the polymerase ORF (Schlicht *et al.*, 1989). It is possible that the HBV DNA polymerase is produced from the 3.5 kb genomic RNA in a similar manner (Chang *et al.*, 1989b).

The replication strategy of HBV requires that a 3.5 kb RNA serve the function of the replication template in the viral life cycle in addition to encoding the nucleocapsid proteins and the HBV DNA polymerase (Will *et al.*, 1987) (Fig. V. 6). It has not been demonstrated for HBV, but by analogy with GSHV, the shortest of the 3.5 kb HBV RNAs is probably the only species which serves as the template for viral replication (Enders *et al.*, 1987). This RNA is encapsidated by HBcAg/P21 polypeptides to form immature core particles which also contain the product of the HBV DNA polymerase open reading frame (Radziwill *et al.*, 1990; Hirsch *et al.*, 1990). The reverse transcriptase and RNase H activities of the HBV DNA polymerase synthesize the long (or minus) strand of HBV genomic DNA using the terminal protein as primer and the pregenomic RNA as template (Miller *et al.*, 1984). The terminal protein is probably encoded in the amino-terminal region of the HBV DNA polymerase open reading frame in a manner similar to that reported for the DHBV terminal protein (Bosch *et al.*, 1988; Bartenschlager & Schaller, 1988). This replication strategy may ensure that both the primer and polymerase required for viral replication are packaged together into the nucleocapsid. Initiation of long strand synthesis probably occurs at the 3' DR1 sequence present in the terminal redundancy of the 3.5 kb RNA (Fig. V. 6), as has been reported for replication of the woodchuck hepatitis virus minus strand DNA (Seeger & Maragos, 1990). Extension of long strand synthesis to the 5' end of the 3.5 kb RNA generates a short terminal redundancy of 8 nucleotides in this strand (Will *et al.*, 1987). The concomitant RNase H degradation of the transcript results in the generation of the 17 nucleotide capped oligoribonucleotide primer from the 5' end of this RNA necessary for initiation of the short (or plus) strand synthesis (Lien *et al.*, 1986; Will *et al.*, 1987; Radziwill *et al.*, 1990). This oligoribonucleotide is translocated by an unknown mechanism from the end of the newly synthesized long strand to the DR2 sequence located 226 nucleotides from the 5' end of the long strand. It serves as the primer for

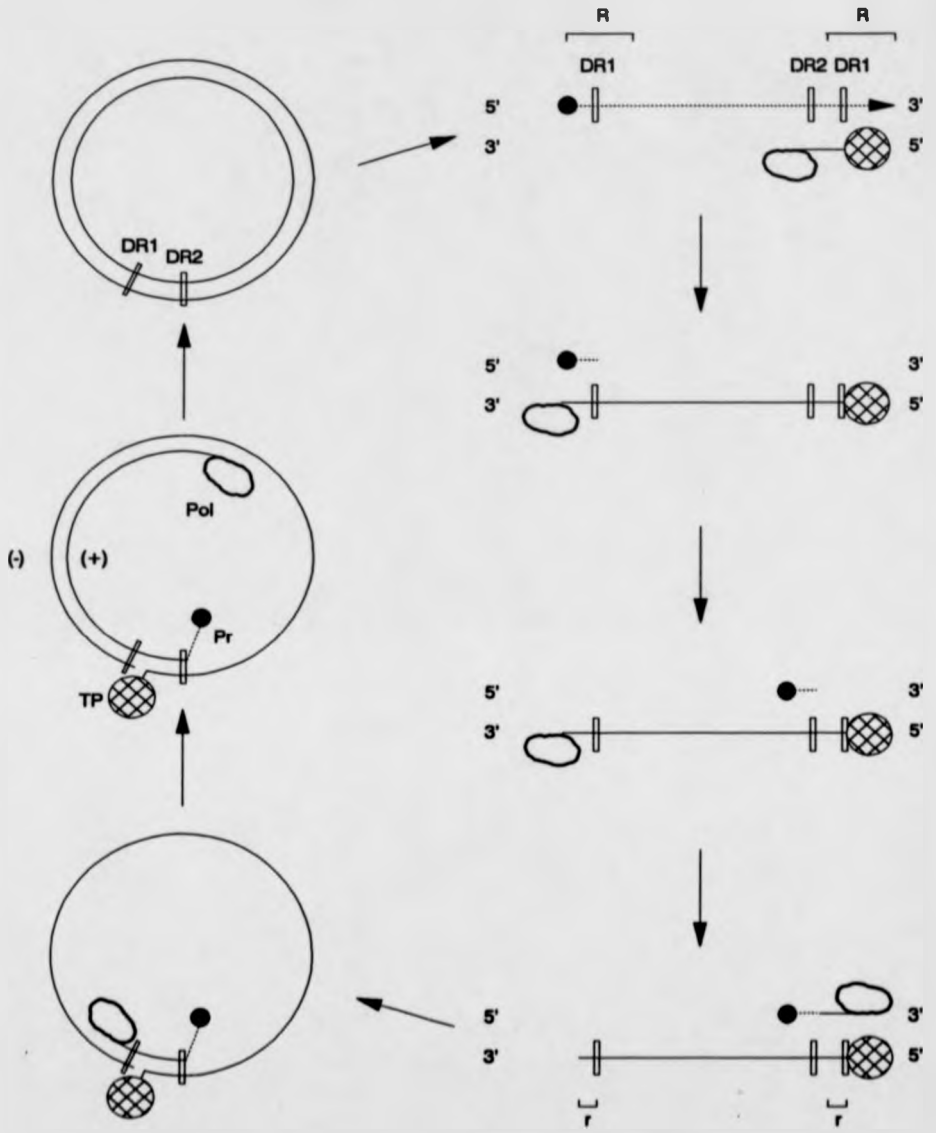


Figure V.6. DNA and RNA intermediates in the replication cycle of HBV. DNA strands are indicated by solid lines and RNA is represented by a dashed line. R, terminal redundancy in the pregenomic RNA; r, terminal redundancy in the HBV minus strand; (+) and (-) indicate the plus (short) and minus (long) strands of the HBV genomic DNA; DR1 and DR2, direct repeats 1 and 2; Pol, HBV DNA polymerase; TP, terminal protein; Pr, oligoribonucleotide primer of plus-strand synthesis.

the short strand synthesis which proceeds from the primer at DR2 to the end of the minus strand (Will *et al.*, 1987). The 8 nucleotide redundancy in the long strand is presumed to permit the transfer of the 3' end of the growing short strand from the protein-linked 5' end to the 3' end of the long strand. This would generate a circular genome structure which upon further extension of the short strand would generate the partially double-stranded HBV genome found in the Dane particle (Robinson *et al.*, 1974; Hruska *et al.*, 1977; Summers *et al.*, 1975; Landers *et al.*, 1977; Sattler & Robinson, 1979; Will *et al.*, 1987) (Fig. V. 1 and 6). The conversion of the pregenomic RNA in immature core particles to the partially double-stranded DNA found in the mature core particles probably occurs in the cytoplasm (Burrell *et al.*, 1982; Gowans *et al.*, 1983; Blum *et al.*, 1984; Gowans *et al.*, 1985; Arakaki *et al.*, 1988; Miller & Robinson, 1984) in a manner similar to that described for DHBV (Summers & Mason, 1982).

The final step of the assembly of the virus particle involves the association of the mature core particle with an appropriately arranged assembly of envelope antigen molecules within the endoplasmic reticulum membrane. The surface antigen assembly subsequently buds into the lumen of the endoplasmic reticulum enveloping the mature core particle (Stein *et al.*, 1972; Huang & Groh, 1973; Yamada & Nakane, 1977; Yamada *et al.*, 1982; Gerber *et al.*, 1988; Kamimura *et al.*, 1981). The complete virion is secreted from the cell through the endoplasmic reticulum and Golgi apparatus pathway (Fig. V. 7). The carbohydrate attached to the surface antigen of the virion and the subviral particles is added during their passage through this pathway (Patzer *et al.*, 1984; Patzer *et al.*, 1986).

The description of the replication cycle of HBV would suggest that the number of covalently closed circular HBV genomes present in the nucleus of an infected hepatocyte would be equal to the number of virus particles which infect that particular cell. Amplification of virus production would therefore depend on each nuclear covalently closed circular HBV DNA molecule producing multiple copies of the pregenomic RNA, which is subsequently reverse transcribed to produce new HBV genomes that will be released from the cell in virus particles. Therefore, one mechanism

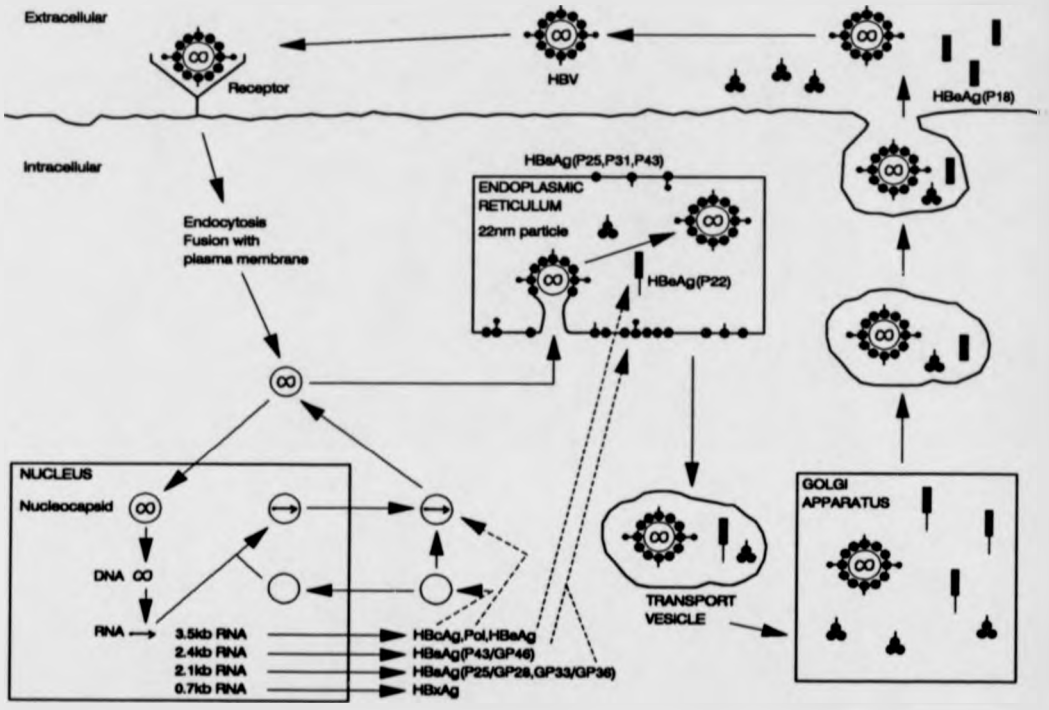


Figure V.7. Intracellular pathway for the synthesis and secretion of HBV, HBsAg subviral particles, and HBeAg polypeptides. The viral components are labelled within the figure. HBsAg, hepatitis B surface antigen; HBcAg, hepatitis B core antigen; HBeAg, hepatitis B e antigen; Pol, HBV DNA polymerase; HBxAg, the product of the X gene.

employed by HBV to amplify virus production is an intrinsic property of the viral replication cycle and depends on the cellular RNA polymerase II transcribing multiple copies of the pregenomic RNA before the HBV template is destroyed in the hepatocyte nuclei.

However, a second mechanism for the amplification of HBV production, similar to that described for DHBV (Tuttleman *et al.*, 1986), may also be operative (Fig. V. 7). Based on the known properties of the HBV polypeptides and the subcellular compartmentalization of the HBV viral and subviral particles, a speculative model for the process leading to the amplification of the nuclear covalently closed circular HBV genomic sequences has been proposed (Raney & McLachlan, 1991; Tuttleman *et al.*, 1986). The mature nucleocapsids that are enveloped by the surface antigen are secreted from the cell as virions. It is possible that this is not the only pathway that mature nucleocapsids follow in the infected cell. In the absence of the envelope antigen, the mature nucleocapsid is probably indistinguishable from the nucleocapsid released from the virion as it infects the hepatocyte. As the fate of this incoming nucleocapsid is probably to carry the HBV genome to the nucleus, it is likely that a newly synthesized mature nucleocapsid which fails to interact with the envelope antigen could return to the nucleus of the cell in which it was synthesized. This could lead to the amplification of the nuclear pool of HBV DNA within a particular cell. The mature nucleocapsid may return to the nucleus of the cell to increase the covalently closed circular HBV DNA pool which is used to generate HBV RNAs, or it may interact with the envelope antigen on the endoplasmic reticulum to produce virions that are secreted from the cell and enter the circulation where they can infect additional hepatocytes. The regulation of the relative frequency with which mature core particles pass through these two pathways will determine the extent of the amplification of the nuclear HBV DNA pool and the level of virus secretion from a particular cell. The relative production of nucleocapsid and envelope antigen may be involved in determining which pathway predominates in the hepatocyte at different times during the life cycle of the virus.

HBV Host Range and Tissue Tropism

In addition to its restricted tissue specificity, HBV has a very limited host range (Gust *et al.*, 1986). HBV has been shown to infect man and chimpanzees (Vaudin *et al.*, 1988; Lichter, 1969; Maynard *et al.*, 1971; Hirschman *et al.*, 1969; Barker *et al.*, 1975). Experimental transmission of HBV to chimpanzees has also been demonstrated reproducibly (Acs *et al.*, 1987; Sureau *et al.*, 1988; Barker *et al.*, 1973; Markenson *et al.*, 1975; Maynard *et al.*, 1972). In addition, there have been reports of experimental transmission of HBV to gibbons (Bancroft *et al.*, 1977), rhesus monkeys (Barker *et al.*, 1975; London *et al.*, 1972), woolly monkeys (Barker *et al.*, 1975) and African green monkeys (London *et al.*, 1972; London *et al.*, 1970). Experimental transmission to additional primate and non-primate species has been unsuccessful (Barker *et al.*, 1975).

The tissue tropism of HBV has been examined in man, chimpanzees and transgenic mice. The primary site of virus replication is the hepatocyte, although evidence of viral infection and replication has been observed in other tissues. HBV nucleic acids and proteins have been observed in various nonhepatic tissues including pancreas, skin, kidney, bile duct epithelium and peripheral blood mononuclear cells (PBMC), with the most extensive studies performed on PBMC (Lamelin & Trepo, 1990; Dejean *et al.*, 1984; Shimoda *et al.*, 1981; Blum *et al.*, 1983). In HBV-infected chimpanzees, viral DNA and RNA were detected in PBMC, including both T and B cells (Korba *et al.*, 1986). The DNA was present primarily in monomeric and multimeric episomal forms, although integrated forms of the DNA were observed. The pregenomic and greater than pregenomic length viral RNAs observed in the chimpanzee PBMC were present at levels approximately 30-fold lower than those seen in the hepatocytes (Korba *et al.*, 1986). Viral antigens have also been detected in human PBMC (Colucci *et al.*, 1986; Zoulim *et al.*, 1988; Hadchouel *et al.*, 1988; Chong-Jin & Jeak-Ling, 1984; Parvaz *et al.*, 1987), the tumor cells of human primary testicular cancer (Somlo, 1989), in human pancreas tissue (Shimoda *et al.*, 1981), and bone marrow aspirates (Elfassi *et al.*, 1984; Romet-Lemonne *et al.*, 1983b; Romet-Lemonne *et al.*, 1983a). The significance of these observations is presently unclear.

The mechanisms responsible for the highly restricted species and tissue tropism of HBV are currently unknown. However, replication of HBV may be restricted to

hepatocytes for several reasons, including the absence of viral receptors on non-hepatic cells and the requirement for liver specific transcription factors for expression from some of the HBV promoters. Characterization of the cell type specificity of the HBV enhancer and promoters has suggested that the enhancer and the core and large surface antigen promoters may direct preferential gene expression in hepatoma cell lines (Jameel & Siddiqui, 1986; Karpen *et al.*, 1988; Antonucci & Rutter, 1989; Honigwachs *et al.*, 1989; Yee, 1989; Raney *et al.*, 1990; Chang *et al.*, 1989), whereas the major surface antigen gene appears to be expressed efficiently in a range of tissue culture cell types (Pourcel *et al.*, 1982; Siddiqui *et al.*, 1986; De-Medina *et al.*, 1988). It has been demonstrated that production of virus particles can be achieved by transfection of hepatoma cells with HBV genomic DNA (Sureau *et al.*, 1986; Tsurimoto *et al.*, 1987; Sells *et al.*, 1987; Chang *et al.*, 1987; Yaginuma *et al.*, 1987). The requirement for hepatoma cells in the production of HBV and WHV by this approach can be circumvented by expressing the pregenomic RNA from an exogenous promoter rather than the core promoter (Junker *et al.*, 1987; Seeger *et al.*, 1989). These observations suggest that in transfection experiments a block to virus production in non-hepatic cells is at the level of transcription from the core promoter. In addition, the large surface antigen promoter has been shown to display liver-cell-type specificity. This promoter contains a sequence element that is necessary for maximal transcriptional activity (Raney *et al.*, 1990; Chang *et al.*, 1989; Nakao *et al.*, 1989) which binds the highly liver-enriched transcription factor, hepatocyte nuclear factor 1 (Courtois *et al.*, 1988; Lichtsteiner & Schibler, 1989; Frain *et al.*, 1989; Chang *et al.*, 1989), and probably explains the preferential expression from this promoter in hepatoma cell lines. In transgenic mice synthesizing HBV particles, it has been shown that the major HBV RNAs are predominantly synthesized in liver and kidney tissue. This indicates that there is a tissue-specific transcriptional restriction in the expression of the HBV genes in this system (Araki *et al.*, 1989; Farza *et al.*, 1988). It also suggests that the block in the HBV life cycle in the murine system is at the level of virus entry into the hepatocyte, possibly indicating the absence of an appropriate receptor in this species. These data suggest that a major factor contributing

to the hepatotropism of HBV may be its highly tissue-specific transcriptional regulation, and that the species host range restriction may be at the level of the viral receptor.

V. B. Hepatitis B Virus Transcriptional Regulation

The characterization of hepadnavirus RNAs in infected liver and various tissue culture systems, and the availability of cloned viral DNA has permitted the analysis of the sequence elements regulating viral transcription. Several viral transcriptional regulatory elements have been identified using DNA mediated gene transfer studies in cell culture (Pourcel *et al.*, 1982; Standring *et al.*, 1984; Shaul *et al.*, 1986; Shaul *et al.*, 1985; Siddiqui *et al.*, 1986; Siddiqui *et al.*, 1987; Treinin & Laub, 1987; De-Medina *et al.*, 1988; Karpen *et al.*, 1988; Raney *et al.*, 1989; Raney *et al.*, 1990; Raney *et al.*, 1991a; Raney *et al.*, 1991b; Honigwachs *et al.*, 1989; Yee, 1989; Chang *et al.*, 1989; Nakao *et al.*, 1989; López-Cabrera *et al.*, 1990; Waisman *et al.*, 1990; Yaginuma & Koike, 1989). Promoter activity has been demonstrated upstream of the HBV large surface antigen, major surface antigen, nucleocapsid and X open reading frames (Fig. V. 2) using transient transfection systems with HBV sequences regulating the expression of reporter genes (Shaul *et al.*, 1985; Shaul *et al.*, 1986; Siddiqui *et al.*, 1986; Siddiqui *et al.*, 1987; Treinin & Laub, 1987; Chang *et al.*, 1989; Raney *et al.*, 1989; Raney *et al.*, 1990; Raney *et al.*, 1991a; Raney *et al.*, 1991b). Detailed characterization of the surface antigen gene regulatory elements has been performed (Shaul *et al.*, 1986; Raney *et al.*, 1989; Raney *et al.*, 1991b; Raney *et al.*, 1992). Seven regions located within 200 nucleotides of the major surface antigen gene transcription initiation site have been identified which influence the level of transcription from this promoter (Raney *et al.*, 1991b). These regions include NF-1 and Sp1 binding sites. The large surface antigen gene regulatory elements were not well characterized until the studies described in this thesis and those of other groups were performed. The identification of important regulatory elements in this promoter region has now been reported (Raney *et al.*, 1991a; Raney *et al.*, 1990; Chang *et al.*, 1989; Zhou & Yen, 1991). The nucleocapsid promoter region has also been

characterized in transfection studies (López-Cabrera *et al.*, 1990; Yaginuma & Koike, 1989; Yee, 1989; Karpen *et al.*, 1988; Honigwachs *et al.*, 1989; Zhang *et al.*, 1992; Zhang *et al.*, 1993; Yuh *et al.*, 1992). This promoter appears to have higher relative activity in differentiated hepatoma cell lines than in dedifferentiated hepatoma or non-hepatoma cell lines, and it contains functional binding sites for the transcription factors C/EBP and Sp1. A minimal promoter element upstream of the X gene has been identified and appears to be modulated by enhancer I, which is adjacent to, or overlaps, the X promoter region (Treinin & Laub, 1987; Zhang *et al.*, 1992; Guo *et al.*, 1991). Two HBV enhancer elements, now referred to as enhancer I and enhancer II, are located between the 3' end of the surface ORF and the 5' end of the X ORF, and in the core promoter region, 5' to the nucleocapsid ORF, respectively (Shaul *et al.*, 1985; Jameel & Siddiqui, 1986; Antonucci & Rutter, 1989; Bulla & Siddiqui, 1988; Faktor *et al.*, 1988; Chang *et al.*, 1987; Wang *et al.*, 1990; Yuh & Ting, 1990; Yuh & Ting, 1991; Vannice & Levinson, 1988; Tognoni *et al.*, 1985; Elfassi, 1987; Zhou & Yen, 1990; López-Cabrera *et al.*, 1991; Guo *et al.*, 1991; Yuh & Ting, 1993). Both enhancer elements display liver-specific characteristics, such as higher levels of transcriptional stimulation in differentiated hepatoma cell lines, DNase I protection using differentiated hepatoma cell line nuclear extracts, and the presence of binding sites for the liver-enriched transcription factor C/EBP (Shaul & Ben Levy, 1987; Patel *et al.*, 1989; Yuh & Ting, 1993; Shaul *et al.*, 1985; Tognoni *et al.*, 1985; Yee, 1989). In addition, a glucocorticoid responsive element has been shown to stimulate surface antigen gene expression, possibly in conjunction with the enhancer I (Tur Kaspa *et al.*, 1988; Tur Kaspa *et al.*, 1986).

The *cis*-acting elements and *trans*-acting factors that are important in the regulation of HBV transcription are parts of a complex system utilized by the virus to facilitate transcription of the appropriate levels of the viral mRNAs. HBV transcription is carried out by the cellular RNA polymerase II. Many of the components of the basal transcription machinery necessary for RNA polymerase II transcription have been identified, and the mechanisms of action by additional transcription factors are currently being elucidated.

V. C. Transcription by RNA Polymerase II

Eukaryotic transcription of messenger RNA by the enzyme RNA polymerase II requires sequence elements and protein factors to direct the enzyme to the correct initiation site on the DNA template in the proper orientation. Typically, the promoter region immediately upstream of the start of transcription contains *cis*-acting sequence elements which bind proteins to influence this process. Many of these promoter regions contain binding sites for transcriptional activators in addition to binding sites for components of the basic transcription machinery. Many promoters contain a TATA box element which specifies the site of initiation and helps to determine the direction of transcription (Breathnach & Chambon, 1981), and an initiator element, Inr, at the initiation site, which also appears to determine the start site (Smale & Baltimore, 1989), especially in the absence of a TATA box element. In addition, enhancer elements, regions of DNA sequence that influence promoter strength in an orientation- and relatively distance-independent manner, are often involved in the regulation of transcription of RNA polymerase II templates.

The process of transcription initiation appears to require an ordered series of events which leads to the regulated expression of genes. In prokaryotes, this process is achieved by a DNA-dependent RNA polymerase in association with different sigma factors which determine promoter specificity. In eukaryotes, three different DNA-dependent RNA polymerases synthesize RNA from distinct sets of genes. RNA polymerase I transcribes ribosomal RNA; RNA polymerase II synthesizes transcripts from protein-encoding genes and most small nuclear RNAs (snRNAs); RNA polymerase III synthesizes small RNAs, such as tRNA and 5S ribosomal RNA. Specific transcription initiation by the eukaryotic RNA polymerases requires additional factors or accessory proteins. The accessory proteins associated with the activity of RNA polymerase II are called general transcription factors, or GTFs. The components required for the formation of a specific, transcription-competent complex have been identified by

determining which chromatographic fractions of cell line or tissue extracts contribute to transcription in *in vitro* transcription systems.

Transcription by RNA polymerase II involves at least seven general transcription factors (GTFs), six of which are essential (Zawel & Reinberg, 1993; Weis & Reinberg, 1992). These factors are called TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH, and TFIIJ. The first step in the assembly of a stable preinitiation complex is the binding of the TFIID component to the promoter approximately 30 nucleotides (-30) upstream of the initiation site (+1) at the AT-rich TATA box element (Buratowski *et al.*, 1989; Ham *et al.*, 1992; Weis & Reinberg, 1992). The TFIID complex is the only GTF shown to have sequence specific promoter binding activity. In the absence of a TATA box element, components of the TFIID fraction can bind either directly or indirectly to the initiator (Inr) or other sequence-specific transcription factors to initiate transcription (Smale *et al.*, 1990). The TFIID complex includes a polypeptide which binds the TATA box element, the TATA binding protein (TBP), and TBP-associated factors (TAFs) (Peterson *et al.*, 1990; Kao *et al.*, 1990). The TAFs appear to be involved in modulating transcriptional activation. They are not necessary for basal levels of transcription in an RNA polymerase II *in vitro* transcription system, but are necessary for activated transcription.

Binding of TFIID to the promoter is the first step in the assembly of the preinitiation complex. This step is stimulated by the association of TFIIA with TFIID. Formation of the TFIID-promoter complex allows the binding of TFIIB to the complex (Maldonado *et al.*, 1990; Buratowski *et al.*, 1989). The TFII(DAB) complex is recognized by RNA polymerase II in association with TFIIF (Killeen *et al.*, 1992). The ordered addition of factors TFIIE, TFIIH, and TFIIJ to the complex results in the formation of a complete preinitiation transcription complex analogous to the prokaryotic closed complex. Transcription initiation requires the hydrolysis of ATP and melting of the DNA to form an open complex. The exact mechanism of this process is unknown although it is suggested that certain GTFs are released (Weis & Reinberg, 1992), possibly after catalyzing the hydrolysis of the ATP. Distinct polypeptides have been identified

for all of the TFII fractions with the exception of TFIIJ, which is not yet well characterized. An additional factor, TFII I, has been identified which binds to the Inr element and can promote TBP binding in the absence of TFIIA (Roy *et al.*, 1991). The yeast homologues of TFIIA, TFIIB, TFIID, TFIIE, and TFIIH have also been identified (Feaver *et al.*, 1991; Gileadi *et al.*, 1992; Sayre *et al.*, 1992a; Sayre *et al.*, 1992b; Henry *et al.*, 1992; Tschochner *et al.*, 1992).

In addition to these components of the general transcription machinery, regulation of transcription involves sequence- specific DNA binding proteins, or activators, which modulate transcriptional activity. These factors appear to interact either directly or indirectly with the general transcription factor complex. One class of sequence specific transcription factors that has been identified contains regions of acidic amino acid residues which appear to be necessary for activation of transcription (Ma & Ptashne, 1987; Triezenberg *et al.*, 1988a; Triezenberg *et al.*, 1988b; Giniger & Ptashne, 1987). These acidic activation domains are found in such transcription factors as the yeast GAL4 and GCN4 activators and the herpes simplex virus VP16 protein (Hope & Struhl, 1986; Ma & Ptashne, 1987; Triezenberg *et al.*, 1988a). Other types of activation domains thus far identified are the glutamine-rich domains of Sp1 and Oct2, and the proline-rich activation domains of CTF/NF-1 and AP-2 (Kadonaga *et al.*, 1988; Courey & Tjian, 1988; Mermod *et al.*, 1989; Mitchell & Tjian, 1989; Williams & Tjian, 1991). The means by which activators communicate with the general transcription machinery have not been fully elucidated, but it is likely that one mechanism of activation involves interaction with TBP, the only specific binding protein of the general transcription machinery (Lin *et al.*, 1991). As the first step in the pathway of assembling a preinitiation complex is thought to be the binding of TBP to the TATA element, TBP is a likely candidate for interaction with other transcription factors which modulate the rate of transcription (Ptashne & Gann, 1990; Stringer *et al.*, 1990). Some activators, such as GAL4 and Sp1, appear to have both an activation domain and a DNA binding domain which allow the activator to bind the upstream sequence and to activate transcription by interaction, either directly or indirectly, with a target among the GTFs (Ptashne &

Gann, 1990). Others, such as VP16 and E1a, lack a DNA binding domain and presumably must act through a DNA binding protein, such as Oct-1 and ATF-2, respectively (Stern & Herr, 1991; Liu & Green, 1990), to affect transcription. Activators may function by recruiting the general transcription factors to the target genes or may change the conformation of the GTFs such that they become active in the process of transcription initiation. Experiments using either the TFIID chromatographic fraction or recombinant TBP in *in vitro* transcription systems suggests that additional factors are required for activation by some transcriptional activators. For example, recombinant TBP is able to support basal transcription but is unable to support activated transcription by such activators as Sp1, CTF/NF-1, and GAL4-VP16, whereas the TFIID fraction can support activated transcription, suggesting that TBP is not the only component of the TFIID fraction required for activated transcription (Tanese *et al.*, 1991; Peterson *et al.*, 1990; Pugh & Tjian, 1990). Additional chromatographic fractionation experiments of *in vitro* transcription extracts and transfection experiments, as well as experiments which demonstrated squelching effects provide evidence for the existence of coactivators, or adaptors, factors which may or may not be closely associated with TBP and are necessary for some activators, such as Sp1 and CTF/NF-1, to mediate their influence. Studies with the human and *Drosophila* coactivator fractions showed that the coactivator fractions were not able to restore activated transcription in response to the transcription factor *zen*, but were able to restore activation by Sp1 and NTF-1, suggesting that there may be multiple coactivators which show species specificity (Hoey *et al.*, 1990; Smale *et al.*, 1990; Dynlacht *et al.*, 1991). The existence of specific coactivators in different tissues could result in more flexibility of transcriptional control. It is possible that activators with different types of activating domains will interact with different coactivators or GTFs. Both *in vivo* and *in vitro* experiments have demonstrated that high concentrations of specific transcription factors can inhibit, or squelch, expression from some promoters lacking the binding site for that transcription factor (Kelleher *et al.*, 1990; Berger *et al.*, 1990; Triezenberg *et al.*, 1988a; Gill & Ptashne, 1988; Meyer *et al.*, 1989). *In vitro* experiments have been performed using a GAL4-VP16 fusion protein

which inhibited transcription from a heterologous promoter. The inhibition was relieved when partially purified TFIID fraction, presumably containing an appropriate coactivator, was added. Inhibition was not relieved when RNA polymerase II or other GTFs were added (Berger *et al.*, 1990; Kelleher *et al.*, 1990). Greater flexibility of transcriptional control could be achieved by increasing the diversity of protein-protein interactions. For example, two activators with the same DNA binding specificities could interact with different coactivators to mediate different functions. Oct-1 and Oct-2, members of the POU domain family of transcription factors, have identical DNA binding specificities, but only Oct-1 binds VP16 to activate transcription during herpes simplex virus infection (Stern *et al.*, 1989; Cleary *et al.*, 1993). The requirement for an additional factor involved in activated transcription by Sp1 on a TATA-less promoter was suggested from experiments in which additional factors present in a human TFIID fraction but not present in *Drosophila* or yeast TFIID fractions were needed for *in vitro*, Sp1 activated transcription of a TATA-less, GC box-containing promoter (Pugh & Tjian, 1990; Smale *et al.*, 1990; Pugh & Tjian, 1991). It was proposed that this factor acted as a "tether" to anchor the preinitiation complex to the promoter, as a TATA box element appears to do.

The regulation of transcription involves a large array of DNA sequence elements and transactivating proteins to transmit the cellular signals to the general transcription machinery. Identifying the types of factors involved, their functional domains and mechanisms of action are important for our understanding of the regulation of differential gene expression. Many regulatory proteins have been grouped into families based on their functional domains and their roles within the cell or organism. One such family of proteins, the homeobox protein family, was first characterized as developmental regulators. This group of proteins contains a characteristic DNA binding domain and includes subfamilies with variations in the DNA binding region of the polypeptide. Many of these proteins are involved in transcriptional regulation.

V. D. Homeodomain Proteins

Homeodomain proteins are a class of proteins involved in developmental regulation in *Drosophila* and higher organisms (Bender *et al.*, 1983; Carrasco *et al.*, 1984; McGinnis *et al.*, 1984a). First identified by the analysis of the *Drosophila Antennapedia* gene (Bender *et al.*, 1983; McGinnis *et al.*, 1984b; Scott *et al.*, 1993), homeotic genes are characterized by a DNA segment of approximately 180 bp long, called the homeo box, which encodes the region of the polypeptide necessary to mediate specific DNA binding. In *Drosophila*, these genes determine the identity and order of body segments. The homeo box regions appear to be highly conserved during evolution and share an open reading frame which encodes the homeo domain of homeotic proteins. Comparison of the protein sequence of this domain revealed homology to the yeast mating-type proteins Mat a1 and Mat α2 which are sequence-specific DNA binding proteins known to control cellular differentiation into the three cell types of yeast (Shepherd *et al.*, 1984; McKay & Steitz, 1981; Johnson & Herskowitz, 1985). The homology between the homeotic genes and the Mat genes suggested a similar functional role in gene regulation, and, consistent with that is the evidence from antibody localization studies that homeo domain-containing proteins accumulate in the nucleus (White & Wilcox, 1984). In addition, many of the homeodomain proteins subsequently identified are sequence-specific transcription factors (Herr *et al.*, 1988; Bodner *et al.*, 1988; Rosenfeld, 1991).

The classic 60 amino acid homeodomain mediates sequence-specific DNA binding by homeodomain proteins. Sequence comparison with prokaryotic and yeast transcriptional regulatory proteins suggested that the homeodomain resembles the helix-turn-helix structure of some DNA binding proteins (Laughon & Scott, 1984; Shepherd *et al.*, 1984; Scott *et al.*, 1993). Determination of the three-dimensional structures of the *Antennapedia*, *engrailed*, and Mat α2 homeodomains has verified the helix-turn-helix structure (Kissinger *et al.*, 1990; Otting *et al.*, 1990; Wolberger *et al.*, 1991). The structure of the homeodomain is a triple α helix with a flexible amino-terminal arm. The second and third helices form the helix-turn-helix motif, with the third helix making contact with the nucleotide bases in the major groove of the DNA. The amino-

terminal arm binds the DNA in the adjacent minor groove, and side chains contact the sugar-phosphate backbone. Amino acid sequence comparison of homeodomains from 38 proteins identified nine invariant amino acids (Gehring, 1987). The *Antennapedia* gene sequence represents the consensus of these sequences, containing 57 of 60 residues most frequently found in each position. The most conserved region was found in the third, recognition, helix and in the amino acids immediately carboxyl-terminal to this helix. This finding was in contrast to the prokaryotic regulatory proteins, which exhibited much more variability in their recognition helices, possibly due to the need to bind to different operator sequences on the DNA. A variant helix-turn-helix homeodomain, the POU domain, has been conserved in several mammalian transcription factors and a yeast developmental regulator. Homeodomain proteins of this subfamily are referred to as POU proteins.

POU Domain Subfamily of Homeodomain Proteins

The cloning of three mammalian transcription factors and a *Caenorhabditis elegans* developmental regulator identified a novel DNA binding domain, the POU domain, which contained a region similar to the homeobox domain of the homeodomain proteins. Like the classic homeodomain proteins, these variant homeodomain-containing proteins, or POU proteins, are involved in the regulation of development (Rosenfeld, 1991; Ruvkun & Finney, 1991; Scholer *et al.*, 1990). The POU domain was first discovered in the pituitary-specific transcription factor Pit-1 (also known as GHF-1), the octamer motif-binding factors Oct-1 and Oct-2, and the product of the nematode cell lineage gene Unc-86, hence the name POU domain (Herr *et al.*, 1988; Bodner *et al.*, 1988; Ingraham *et al.*, 1988; Sturm *et al.*, 1988; Clerc *et al.*, 1988; Finney *et al.*, 1988). Subsequently, other POU proteins have been identified which are involved in cell-type-specific gene regulation, cellular proliferation, DNA replication, developmental control and determination of cell identity (Rosenfeld, 1991; Ruvkun & Finney, 1991; Scholer *et al.*, 1990; Varrijzer *et al.*, 1992). The POU domain is an approximately 160 amino acid region containing two subdomains, the 74 to 82 amino acid amino-terminal POU-specific (POU_s) domain and the 60 amino acid carboxyl-terminal POU-homeodomain

(POU_{HD}) joined by a 15 to 27 amino acid variable linker region (Varrijzer *et al.*, 1992; Assa-Munt *et al.*, 1993; Rosenfeld, 1991). The POU proteins are characterized by these two subdomains, the divergent homeodomain region and the additional POU specific region.

Analysis of the sequence of POU_{HD} predicts that the POU homeodomain is similar to that of the classic homeodomain, containing three α helices, with the third "recognition" helix important for binding the DNA (Ingraham *et al.*, 1990). Unlike the classic homeodomain, POU_{HD} is unable to bind DNA with specificity and high affinity (Ingraham *et al.*, 1990; Sturm & Herr, 1988; Kristie & Sharp, 1990; Verrijzer *et al.*, 1990a; Verrijzer *et al.*, 1990b; Aurora & Herr, 1992). It has been shown that the POU_S domain of Pit-1 is necessary (in addition to POU_{HD}) for site specific, high affinity DNA binding. It has also been shown that Oct-1 binds the octamer motif (Verrijzer *et al.*, 1992; Aurora & Herr, 1992), with the POU_{HD} domain recognizing the 3' end and the POU_S domain recognizing the 5' end of the motif (Verrijzer *et al.*, 1990a). Both of the domains have contacts in the major groove of the DNA (Aurora & Herr, 1992) and binding results in bending of the DNA molecule (Verrijzer *et al.*, 1991). Comparison of several POU protein sequences indicated two regions of very high homology in the POU_S domain, called the POU_S-A and POU_S-B subdomains. The predicted structure of the POU_S domain was that of two α helices, and mutational analysis suggested that the first helix and clusters of basic amino acid residues in the POU_S domain were important for DNA binding (Ingraham *et al.*, 1990). Recently, the solution structure of the Oct-1 POU_S domain was determined and demonstrated that this region consists of four packed α helices surrounding a hydrophobic core (Assa-Munt *et al.*, 1993). This structure is similar to the DNA binding domains of the bacteriophage lambda and 434 repressors and 434 Cro, which are helix-turn-helix motifs. The POU_S domain differs from the conventional helix-turn-helix motif in its extra length in helix two, the first helix of the helix-turn-helix motif, and in the turn, which is the spacer region between the subdomains A and B. A model based on the solution structure of the Oct-1 POU_S domain and the similarity to the bacteriophage repressor and 434 Cro structure suggests

that the third α helix of the POU_S domain is equivalent to the recognition helix of the helix-turn-helix motif, which lies in the major groove of the DNA and makes sequence-specific contacts with the DNA (Assa-Munt *et al.*, 1993). This helix is the most highly conserved region of the POU_S domain, consistent with an important role in sequence-specific DNA binding. In addition, the importance of certain conserved residues, such as the glutamines present at the beginning of helices two and three, are consistent with mutations of the Oct-1 and Pit-1 POU_S domains which affect DNA binding (Sturm & Herr, 1988; Ingraham *et al.*, 1990). It appears that the POU-specific and POU-homeodomain regions of the POU domain are physically linked domains comprising helix-turn-helix motifs, both of which are necessary for high affinity, sequence-specific DNA binding.

V. E. Hepatocyte Nuclear Factor 1

The liver enriched transcription factor hepatocyte nuclear factor 1, HNF1, is a member of the POU subfamily of homeodomain proteins. HNF1, which has also been called LFB1, APF, and HP1 (Cereghini *et al.*, 1988; Hardon *et al.*, 1988; Schorpp *et al.*, 1988), contains a highly divergent homeodomain and a region of homology to the POU_S domain A box. This 88 to 93 kD glycosylated protein was originally purified from rat liver nuclear extracts using a β fibrinogen promoter element (Courtois *et al.*, 1988; Courtois *et al.*, 1987) and is involved in the regulation of many genes expressed predominantly in the liver, including α fetoprotein, albumin, transthyretin, α1-antitrypsin, and the large surface antigen gene of HBV (Mendel & Crabtree, 1991; Hardon *et al.*, 1988; Monaci *et al.*, 1988; Courtois *et al.*, 1988; Feuerman *et al.*, 1989; Chang *et al.*, 1989; Raney *et al.*, 1991a). Cloning of the HNF1 cDNA which encodes the 628 amino acid protein has permitted the characterization of the functional domains of the polypeptide (Fig. V. 8) (Baumhueter *et al.*, 1990; Chouard *et al.*, 1990; Frain *et al.*, 1989; Nicosia *et al.*, 1990; Raney *et al.*, 1991a). The amino-terminal 281 amino acids contain regions important in dimerization and specific DNA binding. The carboxyl

Figure V.8. Organization of the functional domains of the HNF1 polypeptide. DIM, dimerization domain; POU-A, region of POU subdomain A homology; HOMEO, homeoboxlike domain; PRO, proline-rich transcriptional activation domain ; GLN/PRO, glutamine- and proline-rich transcriptional activation domain; SER, serine-rich transcriptional activation domain.

terminal portion of the polypeptide contains three regions identified as transcriptional activation domains (Raney *et al.*, 1991a; Nicosia *et al.*, 1990; Toniatti *et al.*, 1993).

The amino-terminal region necessary for DNA binding comprises three functional domains, the dimerization domain located in the first 32 residues, a domain with homology to the POU_S A subregion, and an unusually long (81 amino acids) divergent homeodomain. The structural requirements for DNA binding appear to be more complex in HNF1 than in other POU proteins, which require the POU_S and POU_{HD} regions, and homeodomain proteins, whose homeodomain alone is sufficient for specific DNA binding (Sturm & Herr, 1988; Muller *et al.*, 1988). HNF1 binds its recognition sequence as a dimer. The predicted structure of the dimerization domain of HNF1 is similar to that of the rod portion of cardiac myosin, which is folded in long α helices wrapped in a two-strand coiled coil with hydrophobic side chains aligned on one side (Nicosia *et al.*, 1990). A peptide containing this domain folds into two α helices kinked to form a parallel homodimer stabilized by the hydrophobic faces (Pastore *et al.*, 1991; De Francesco *et al.*, 1991). A protein factor, called the dimerization cofactor of HNF1 (DCoH) interacts with and stabilizes the HNF1 dimers by reducing the rate of monomer exchange (Mendel *et al.*, 1991b). DCoH also enhances HNF1-mediated transactivation although it has no DNA binding or transactivation capability itself (Mendel *et al.*, 1991b). Mutational analysis of this region has demonstrated that DNA binding, as well as dimerization, is dependent upon the HNF1 dimerization domain (Nicosia *et al.*, 1990; Chouard *et al.*, 1990). It has also been shown that HNF1 can dimerize in solution in the absence of DNA (Chouard *et al.*, 1990). The POU-like domain amino-terminal to the divergent homeodomain has 27% homology with the POU_S A subdomain and appears to be involved in increasing the specificity of the DNA binding by the homeodomain (De Simone & Cortese, 1992), as does the POU_S region of POU proteins. The structure of the long (81 versus 60 amino acid residues) homeodomain of HNF1 does not conform exactly to the helix-turn-helix motif. The alignment of sequences necessary for the best homology calls for "looping out" 21 amino acids between helices two and three (Finney, 1990). This loop would replace the

conventional turn between these two helices in a region proposed to be tolerant of insertions without affecting DNA binding activity. Two alternative loops, of 18 and 20 residues, which would disrupt helix two, have been suggested as likely based on the ability of deletion mutants of this region to bind DNA (Nicosia *et al.*, 1990). The third helix of the HNF1 homeodomain contains the highest homology with the recognition helix of homeodomains and deletion of this helix abolishes DNA binding (Chouard *et al.*, 1990; Nicosia *et al.*, 1990).

Although HNF1 is the first mammalian homeodomain protein identified to bind DNA as a dimer, other transcription factors, such as MyoD and members of the AP-1 family, are known to form dimers prior to DNA binding (Turner & Tjian, 1989; Ransone *et al.*, 1989). The ability of transcription factors to interact with other proteins with similar dimerization faces, to form heterodimers, can create new DNA binding specificities and regulatory activities. The consensus binding site sequence of HNF1, GTTAATNATTAAC, is an inverted palindrome suggestive of dimer binding. Although HNF1 does bind its recognition sequence as a homodimer, it is interesting that none of the identified binding sites contains a perfect inverted repeat (Mendel & Crabtree, 1991; Frain *et al.*, 1989; Nicosia *et al.*, 1990). Most of the binding sites contains a perfect or near perfect half site, with a divergent second half site, which might suggest that other HNF1-like proteins with similar dimerization domains could bind the sequence as a heterodimer with HNF1 and perhaps broaden the functions of HNF1 by altering its DNA binding specificity or activities, depending upon its heterodimer partner.

HNF1 heterodimerizes with a related protein called variant HNF1 (vHNF1 or HNF1β), which binds to the same DNA sequence as HNF1 and was first identified in a dedifferentiated hepatoma cell line (Baumhueter *et al.*, 1988; Cereghini *et al.*, 1988; Rey-Campos *et al.*, 1991). In tissue culture cell lines, HNF1 appears to be expressed only in differentiated hepatoma lines, whereas vHNF1 is expressed in dedifferentiated hepatoma cell lines and extinguished somatic cell hybrids and in very low levels in differentiated hepatoma cell lines (Rey-Campos *et al.*, 1991; Cereghini *et al.*, 1990; Baumhueter *et al.*, 1988). The tissue distribution of HNF1 *in vivo* is broader than was initially expected. It

is found not only in the liver, but also in the kidney, the epithelial cells of the intestine, and in very low levels in stomach, pancreas, spleen and testis (Cereghini *et al.*, 1992; Baumhueter *et al.*, 1990; Kuo *et al.*, 1990; Tronche & Yaniv, 1992; Blumenfeld *et al.*, 1991). vHNF1 expression overlaps considerably with that of HNF1, but the levels of expression in the tissues differs from HNF1. HNF1 mRNA is expressed at much higher levels than vHNF1 transcripts in mouse liver and intestinal tissue, but in comparable amounts in kidney (Rey-Campos *et al.*, 1991). vHNF1 mRNA has been observed in liver, kidney, intestine, stomach, pancreas, ovary, and lung (De Simone *et al.*, 1991; Ott *et al.*, 1991; Tronche & Yaniv, 1992; Bach *et al.*, 1991; Rey-Campos *et al.*, 1991). During development in the mouse, vHNF1 expression precedes that of HNF1 by one to two days (Ott *et al.*, 1991; Cereghini *et al.*, 1992). Isolation of the cDNA for vHNF1 from dedifferentiated cells allowed the characterization of the 557 amino acid vHNF1 polypeptide structure (Rey-Campos *et al.*, 1991). The amino acid sequences of vHNF1 and HNF1 are highly conserved in the amino-terminal region of the polypeptide, with 72% identity in the dimerization domain, 66% identity in the region of POU homology, and 92% identity in the homeodomain (Mendel *et al.*, 1991a; Rey-Campos *et al.*, 1991). The activation domains identified by *in vitro* transcription analysis in the carboxyl-terminal portion of the polypeptide are not conserved in the vHNF1 polypeptide (Nicosia *et al.*, 1990; Rey-Campos *et al.*, 1991; De Simone *et al.*, 1991; Mendel *et al.*, 1991a). In fact, there is little conservation in the carboxyl-terminal portion of the molecule, with the exception of the region of the polypeptide containing the glutamine- and proline-rich activation domain identified in transient transfection experiments (Raney *et al.*, 1991a). The broad region containing the transactivation domain displays 54% amino acid identity (71 of 132 residues) and the 18 amino acid region most critical for transactivation of the HBV large surface antigen by the HNF1 polypeptide has 78% amino acid identity (Raney *et al.*, 1991a). The vHNF1 polypeptide can transactivate the albumin promoter in transient transfection studies (Rey-Campos *et al.*, 1991). It is possible that this transactivation is mediated by the conserved regions of the HNF1 and vHNF1 polypeptides which were identified as transcriptional activation domains using

the HBV large surface antigen promoter and human albumin HNF1 binding site sequences in transfection studies (Raney *et al.*, 1991a). The coexpression of vHNF1 and HNF1 in certain tissues and some cell lines, and their ability to form heterodimers and to recognize the same DNA binding site suggests a possible role for differential transcriptional regulation by the formation of vHNF1/HNF1 heterodimers. The relative amounts of these polypeptides in various tissues may be important factors in the tissue-specific regulation of HNF1-recognition-site-containing promoters.

It is most likely that tissue-specific transcriptional regulation is the result of the combined actions of many factors, both ubiquitous and tissue-specific. Several elements control the liver-specific expression of the albumin gene, including the ubiquitous factors NFY and NF-1 as well as the liver-enriched C/EBP and HNF1 (Maire *et al.*, 1989; Lichtsteiner & Schibler, 1989; Wuarin *et al.*, 1990). Other liver-specific genes, such as α1-antitrypsin, contain DNA elements that bind multiple liver-enriched factors such as HNF4 and HNF1 (Tronche & Yaniv, 1992). Negative regulation also appears to have a role in tissue-specific expression, as expression of α fetoprotein (AFP) is turned off in hepatocytes just after birth. The loss of AFP expression is accompanied by the loss of *in vivo* DNase I hypersensitivity of the promoter, and possibly by displacement of HNF1, which is a major positive regulator of AFP expression (Vacher & Tilghman, 1990; Deng & Karin, 1993). The broad range of expression of HNF1 also argues that this factor alone is probably not responsible for the liver-specific transcription of genes containing HNF1 binding sites in their promoters. In this regard, it is worth noting that the networks of factors involved in differential gene expression appear to regulate themselves. The control of expression of the HNF1 gene includes positive regulation by the factors HNF4 and HNF3. Mutation of the binding sites for these factors in the HNF1 promoter reduces its activity (Kuo *et al.*, 1992; Tian & Schibler, 1991). It has also been reported, though not demonstrated in the literature, that HNF1 negatively regulates itself, suggesting the importance of maintaining precise concentrations of HNF1 in the cell (Tronche & Yaniv, 1992). Other factors, such as HNF3 and C/EBP, are positively auto-regulated (Tronche & Yaniv, 1992). Therefore, it appears that a

complex, hierarchical network including liver-enriched and ubiquitous factors is necessary to control the transcription of liver-enriched transcription factor genes which are involved in the control of liver development and the maintenance of the liver cell phenotype.

V. F. Objectives

Hepatitis B virus infection is primarily restricted to hepatocytes. Although the exact mechanisms governing this tropism are not fully understood, it is likely that liver-specific transcriptional regulation of HBV occurs. The characterization of cell-type-specific *cis*-acting elements has suggested that the viral enhancer I, the nucleocapsid promoter, and the large surface antigen promoter display preferential activity in differentiated hepatoma cell lines. The distribution of viral mRNAs expressed in transgenic mice which synthesize HBV particles suggests that, in addition to the nucleocapsid and large surface antigen promoters, the major surface antigen promoter is more active in liver tissue than in other tissues. A detailed analysis of the cell-line- or tissue-specific transcriptional regulation of the HBV promoters may elucidate some of the mechanisms of hepatotropism and their role in the biogenesis of the virus.

The major aim of this project was to characterize the transcriptional regulation of the expression of the large surface antigen gene of hepatitis B virus. To achieve this, the objectives of this study were to identify the regions of the HBV genome involved in the transcriptional regulation of the large surface antigen gene, particularly the regulatory elements responsible for liver-cell-type specificity of large surface antigen gene expression, and to identify the specific transcription factors involved in the regulation of the large surface antigen promoter and examine their roles in the liver-cell-type transcriptional activity of the large surface antigen promoter. The characterization of the large surface antigen promoter should lead to a better understanding of the role of transcriptional regulation in the hepatotropism of hepatitis B virus.

VI. MATERIALS AND METHODS

VI. A. Plasmid Constructions

The steps in the cloning of the plasmid constructs used in the transfection experiments were performed by standard techniques (Sambrook *et al.*, 1989). HBV DNA sequences in these constructions were derived from the plasmid pCP10, which contains two copies of the HBV genome (subtype *ayw*) cloned into the *Eco*RI site of pBR322 (Dubois *et al.*, 1980). The firefly luciferase (LUC) reporter gene in these constructions was derived from the plasmid, p19DLUC. p19DLUC was generated from p19LUC (Fig. VI. 1) (Van Zonneveld *et al.*, 1988) by deletion of the polylinker sequences between the *Hind*III and *Sal*I sites and insertion of a *Hind*III linker at this position. This eliminated the polylinker *Sph*I site containing the sequence ATG, which might have interfered with the translation of the luciferase gene product from the RNA transcripts synthesized from the HBV promoters.

The plasmids SpLUC, XpLUC, CpLUC and PS(1)pLUC each contain one complete HBV genome located directly 5' to the promoterless LUC reporter gene such that the expression of the LUC gene is governed by the hepatitis B major surface antigen (Sp), X gene (Xp), nucleocapsid, or core, (Cp), and large surface antigen [PS(1)] promoters, respectively (Fig. VII. 4). These four plasmids were constructed in our laboratory by Alan McLachlan. The plasmid SpLUC was constructed by digesting pCP10 with *Xho*I, filling in the overhang with the Klenow fragment of *E. coli* DNA polymerase, ligating a *Hind*III linker [d(CAAGCTTG)] using *T4* DNA ligase, digesting with *Hind*III and cloning the 3.2-kbp HBV fragment into the *Hind*III site of the plasmid p19DLUC. The unique HBV *Xho*I site used in SpLUC is located 157 nucleotides 3' to a predominant surface antigen gene transcription initiation site located at nucleotide 3159 in the HBV genome (Cattaneo *et al.*, 1983; Raney *et al.*, 1989) (see Fig. V. 2 for genome organization). Therefore, the plasmid SpLUC contains one complete HBV genome

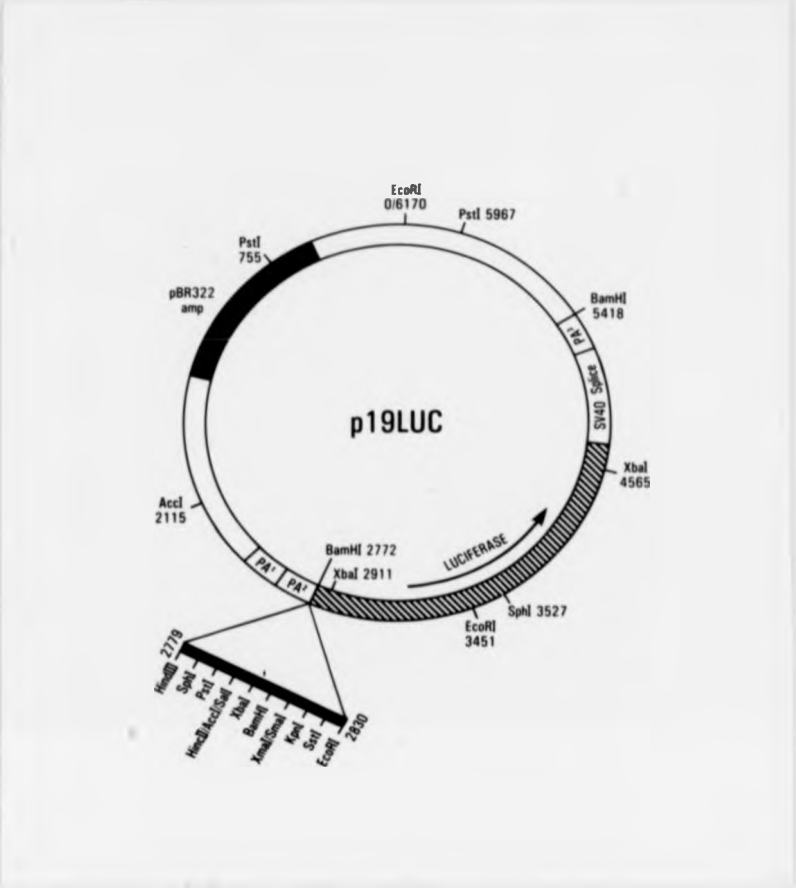


Figure VI.1. **Restriction map of luciferase reporter gene plasmid p19LUC (Van Zonneveld *et al.*, 1988). The vector includes the luciferase ORF followed by an SV40 splice site and polyadenylation site (pA^3). A multiple cloning site lies upstream of the luciferase ORF and is preceded by two SV40-derived polyadenylation sites (pA^1 and pA^2) and the pBR322 ampicillin resistance gene. p19LUC was modified by deletion of the *Sph*I and *Pst*I restriction endonuclease sites to generate the vector p19DLUC for the construction of HBV promoter-luciferase constructs. Several restriction endonuclease sites are indicated.**

(nucleotides 130 to 3182/1 to 133') located directly 5' to the promoterless LUC reporter gene such that the expression of the LUC gene is governed by the hepatitis B major surface antigen promoter. The designation 133' has been used to indicate that the nucleotides 130 to 133 are present twice in this plasmid and that nucleotides 130 to 133 and 130' to 133' are distal and proximal to the LUC ORF, respectively. The plasmid XpLUC was constructed in a manner similar to SpLUC, using the unique *Nco*I site located 67 nucleotides 3' to a predominant X gene transcription initiation site located at nucleotide 1310 in the HBV genome (Siddiqui *et al.*, 1987; Treinin & Laub, 1987) to insert the HBV genome into p19DLUC. XpLUC contains one complete HBV genome (nucleotides 1375 to 3182/1 to 1376') located 5' to the LUC open reading frame (ORF) such that expression of the LUC gene is governed by the HBV X gene promoter. The plasmid CpLUC was constructed similarly using the unique *Fsp*I site located 20 nucleotides 3' to a nucleocapsid gene transcription initiation site located at nucleotide 1785 in the HBV genome (Yaginuma *et al.*, 1987; Sells *et al.*, 1988) to insert the HBV genome into p19DLUC. CpLUC contains one complete HBV genome (nucleotides 1805 to 3182/1 to 1804) located 5' to the LUC ORF such that expression of the LUC gene is governed by the HBV nucleocapsid promoter. The plasmid PS(1)pLUC was constructed by isolating a *Bgl*II fragment from pCP10 containing the nucleotides 2840 to 3182/1 to 2843', and ligating this fragment into the *Bam*HI site of pUC13 (Vieira & Messing, 1982), generating the plasmid pHBV(BH)PS(1)p. The pHBV(BH)PS(1)p plasmid was digested with *Sma*I at the *Sma*I site contained in the pUC13 polylinker segment of this plasmid and a *Hind*III linker [d(CAAGCTTG)] was inserted. The plasmid was then digested with *Hind*III and the 3.2 kbp HBV fragment was inserted at the *Hind*III site of p19DLUC. The *Bgl*II site used for making this construction is located 35 nucleotides 3' to a predominant large surface antigen gene transcription initiation site located at nucleotide 2809 in the HBV genome (Yaginuma *et al.*, 1987; Araki *et al.*, 1989; Sells *et al.*, 1988; Farza *et al.*, 1988). The plasmid PS(1)pLUC contains one complete HBV genome (nucleotides 2840 to 3182/1 to 2843') located 5' to the LUC ORF such that the expression of the LUC gene is governed by the HBV large surface antigen promoter.

The nucleotide sequences are designated by using coordinates derived from the Genbank genetic sequence data bank.

The construction of the PS(1)pLUC 5' deletion series (Fig. VII. 6) was performed by digestion of HBV sequences in pHBV(BH)PS(1)p with appropriate restriction endonucleases or *Bal*31 nuclease and subsequent cloning steps similar to those used to generate PS(1)pLUC. All of the 5' deletion series plasmids contain the HBV sequences from -25 to +35 relative to the transcriptional initiation site at nucleotide 2809 upstream of the luciferase sequences. The plasmid PS(1)pΔ2840-129LUC was made by digestion of pHBV(BH)PS(1)p with *Xho*I and *Sal*I, filling in the overhangs as required, ligating, selecting the deleted plasmid, digesting this with *Hind*III, and gel isolation of the *Hind*III fragment which was then ligated to the *Hind*III site of p19DLUC. The plasmids PS(1)pΔ2840-827LUC, PS(1)pΔ2840-1070LUC, PS(1)pΔ2840-1116LUC, PS(1)pΔ2840-1238LUC, PS(1)pΔ2840-1374LUC, PS(1)pΔ2840-1402LUC, PS(1)pΔ2840-2143LUC, and PS(1)pΔ2840-2425LUC were generated by the same protocol, using *Acc*I, *Nsi*I and *Sal*I, *Stu*I and *Sal*I, *Sph*I and *Sal*I, *Nco*I and *Sal*I, *Bam*HI and *Sal*I, *Xba*I, and *Bgl*II and *Hinc*II, respectively, in the first digestion. The plasmids PS(1)pΔ2840-2547LUC, PS(1)pΔ2840-2612LUC, PS(1)pΔ2840-2707LUC, PS(1)pΔ2840-2718LUC, PS(1)pΔ2840-2733LUC, PS(1)pΔ2840-2734LUC, PS(1)pΔ2840-2767LUC, and PS(1)pΔ2840-2783LUC were generated by digestion of pHBV(SS)PS(1)pLUC with *Bgl*II followed by *Bal*31 nuclease digestion (International Biotechnologies, Inc.), digestion with *Sal*I to delete upstream sequences, filling in the overhangs with the Klenow fragment of *E. coli* polymerase, and ligating. The *Bal*31 nuclease protocol was performed as suggested by the manufacturer. All deletion breakpoints generated by *Bal*31 digestion were determined by dideoxynucleotide sequencing (Sanger *et al.*, 1977) using the *Sequenase 2.0* sequencing kit from United States Biochemical Corporation. The extent of the deleted nucleotide sequence is indicated in the plasmid designation using coordinates from the Genbank genetic sequence data bank. The plasmid pHBV(SS)PS(1)pLUC was used for the generation of the *Bal*31 5' deletion set because it contained a *Sal*I site upstream of the HBV sequences which could be used to delete the upstream sequences

after the *Bal*31 digestion. pHBV(SS)PS(1)pLUC is the same as PS(1)pLUC except in the region of the polylinker. The HBV genome was inserted into p19DLUC as a *Sal*I to *Sac*I fragment from pHBV(BH)PS(1)p and so contains an *Ava*I and a *Sac*I site 3' to the *Hind*III site at the 3' end of the fragment. At the 5' end, the HBV fragment lacks the *Acc*I, *Hinc*II, and *Pst*I sites contained in PS(1)pLUC, but contains the p19DLUC polylinker sequences from *Sal*I to *Hind*III. Because the full-length PS(1)pLUC plasmids differ from each other by a few nucleotides in the region that will be transcribed, both were transfected to ensure that these changes did not affect luciferase activity generated by the plasmids.

The clustered point mutation plasmids PS(1)pM1LUC, PS(1)pM2LUC, PS(1)pM3LUC, PS(1)pM4LUC, and PS(1)pΔ2840-2425M5LUC, were generated by priming polymerase chain reactions (PCR) with mutant oligonucleotides to make the desired mutation within the HBV sequence, and subsequent cloning steps to insert the mutated HBV fragments into the luciferase vector. To generate the four mutant plasmids PS(1)pM1LUC, PS(1)pM2LUC, PS(1)pM3LUC, and PS(1)pM4LUC, a two-step mutant PCR method was employed (Landt *et al.*, 1990). 200 ng of the oligonucleotide LUC29, whose sequence is TGCTCTCCAGCGGTTCC and is contained in the luciferase segment of PS(1)pLUC, and 200 ng of an oligonucleotide containing a mutation of a region of the large surface antigen gene promoter [PS(1)pM1 to M4] were used to prime the initial PCR using 10 ng of the plasmid PS(1)pΔ2840-2707LUC as template (Fig. VI. 2). The same luciferase oligonucleotide (LUC29) was used to prime each of the four reactions. The oligonucleotide PS(1)pM1, used in the first PCR to generate PS(1)pM1LUC, spanned nucleotides 2724 to 2757 and contained a mutation of the sequence between nucleotides 2736 and 2745. The sequence of PS(1)pM1 is ATCATTACTTCCTTGGGTACCCCTATTTACACAC. The oligonucleotide PS(1)pM2, used in the first PCR to generate PS(1)pM2LUC, spanned nucleotides 2734 to 2767 and contained a mutation of the sequence between nucleotides 2746 and 2755. The sequence of PS(1)pM2 is CCAAACTAGACATGGGTACCCTACTCTATGGAAG. The oligonucleotide PS(1)pM3, used in the first PCR to generate PS(1)pM3LUC, spanned

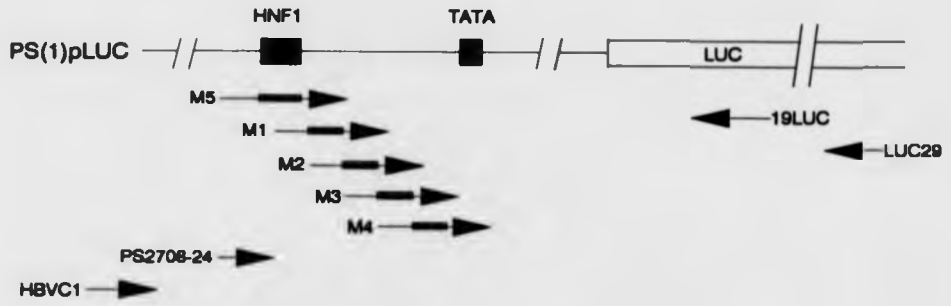


Figure VI.2. Diagrammatic representation of the location of oligonucleotide primers used for synthesis of plasmid mutations by polymerase chain reaction (PCR). The HBV sequences of plasmid PS(1)pLUC are represented by a thin line. The locations of the HNF1 and TATA box sequence homologies are indicated by boxes labelled HNF1 and TATA, respectively. The luciferase gene sequences are indicated by the box labelled LUC. The locations and orientations of the single-stranded oligonucleotide primers 19LUC, LUC29, PS2708-24, HBVC1, M1, M2, M3, M4 and M5 are indicated by arrows. The shaded box on the M1, M2, M3, M4 and M5 arrows represent the mutations introduced into the HBV sequences.

nucleotides 2744 to 2777 and contained a mutation of the sequence between nucleotides 2756 and 2765. The sequence of PS(1)pM3 is CACTATTTACACTTGGGTACCCAGGCGGGTATAT. The oligonucleotide PS(1)pM4, used in the first PCR to generate PS(1)pM4LUC, spanned nucleotides 2754 and 2787 and contained a mutation of the sequence between nucleotides 2766 and 2775. The sequence of PS(1)pM4 is ACACTCTATGGATTAGGTACCCATTATATAAGAG. Each of the four first PCRs generated a single product of the expected size, 249 bp [PS(1)pM1], 239 bp [PS(1)pM2], 229 bp [PS(1)pM3], and 219 bp [PS(1)pM4]. These fragments were phenol extracted to remove protein, ether extracted to remove phenol residue, and ethanol precipitated. The second PCR was primed with 50% of the first product as the luciferase, or downstream, oligonucleotide and 200 ng of an oligonucleotide spanning HBV nucleotides 2708 to 2724 (PS2708-24) as the upstream oligonucleotide. The oligonucleotide PS2708-24 contained seven additional bases at its 5' end to generate a *Sal*I restriction site. The sequence of PS2708-24 is TTGTCGACCAGAACATCTAGTTAA. 25 ng of the template PS(1)pΔ2840-2707LUC were used in the second PCR. The second PCR generated a product of 272 bp, which was digested with the enzymes *Sac*I and *Sal*I, gel isolated, and ligated into p19DLUC at the *Sac*I and *Sal*I sites in the polylinker, generating the plasmids PS(1)pM1LUC, PS(1)pM2LUC, PS(1)pM3LUC, and PS(1)pM4LUC. These plasmids are derivatives of the plasmid PS(1)pΔ2840-2707LUC, each containing a 10 bp mutation in the region between the HNF1 binding site and TATA box element in the large surface antigen promoter. The sequence of the fragments produced by PCR were verified by dideoxynucleotide sequencing. The construct PS(1)pM1LUC lacked the C nucleotide within the mutation at position 2745, which did not alter the mutation within that region, but did shorten the spacing by one nucleotide between the HNF1 and TATA box elements. The construct PS(1)pM4LUC contained a point mutation at nucleotide 2753 (C to T). All other PCR generated fragments contained the expected sequence.

The plasmid PS(1)pΔ2840-2425M5LUC was generated essentially by the same technique. PS(1)pΔ2840-2425M5LUC contains the promoter sequences from -383 to +35

and is designed to be the same as PS(1)pΔ2840-2425LUC with the exception of a 13 bp mutation in the HNF1 binding site from nucleotide 2720 to 2732. This construct was generated by the two-step mutant PCR protocol using 200 ng of the oligonucleotide 19LUC, whose sequence GGCGTCTTCCATTTTACCAACA, is contained in the luciferase segment of PS(1)pLUC, and 200 ng of an oligonucleotide which spanned nucleotides 2708 to 2744 and contained the mutation from nucleotide 2720 to 2732 (PS(1)pM5, whose sequence is CCAGAACATCTATACTGGGTACCTCTCCAAACTAGAC) to prime the first PCR on 10 ng of the template PS(1)pLUC. The 225 bp product from the first PCR was gel isolated and 50% was used as the downstream oligonucleotide for the second PCR. 200 ng of the oligonucleotide HBVC1 (CATGGACATCGACCCTTATA) which spanned HBV nucleotides 1902 to 1921 were used as the upstream oligonucleotide, and 25 ng of the template PS(1)pLUC were used for the second PCR. The 1032 bp product was phenol and ether extracted, ethanol precipitated and digested with *Sac*I and *Bgl*II to generate a 459 bp fragment which was then gel isolated for cloning into Bluescript SK(-) (Stratagene, Inc.). The mutated *Sac*I to *Bgl*II fragment was inserted into the *Bam*HI and *Sac*I sites in the Bluescript polylinker create PS(1)pM5BS. The mutated, PCR synthesized HBV fragment was excised from Bluescript by digestion of PS(1)M5BS with *Hind*III, gel isolated and inserted into the *Hind*III site of the p19DLUC, generating PS(1)pΔ2840-2425M5LUC. The sequences of the fragments generated by PCR were verified by dideoxynucleotide sequencing.

The plasmids PS(1)pΔ2840-2767HNF1(+)LUC, PS(1)pΔ2840-2767HNF1(-)LUC, PS(1)pΔ2840-2767(2)HNF1LUC, PS(1)pΔ2840-2767(10)HNF1LUC, PS(1)pΔ2840-2767AHNF1(+)LUC are derivatives of the HBV PS(1) minimal promoter construct PS(1)pΔ2840-2767LUC in which a double-stranded HNF1 binding site oligonucleotide has been inserted immediately 5' to the HBV sequences. The plasmids PS(1)pΔ2840-2767HNF1(+)LUC and PS(1)pΔ2840-2767HNF1(-)LUC were constructed by inserting a double-stranded oligonucleotide containing the PS(1) promoter HNF1 binding site (spanning HBV oligonucleotides 2719 to 2734), produced by annealing the

oligonucleotides AGCTAGTTAATCATTACTTC and AGCTGAAGTAATGATTAACT, into the unique *Hind*III site in the plasmid PS(1)pΔ2840-2767(H)LUC in the same (+) or opposite (-) orientation to the HNF1 binding site in the large surface antigen promoter. The plasmid PS(1)pΔ2840-2767(H)LUC is a derivative of PS(1)pΔ2840-2767LUC in which the *Hind*III site 3' to the HBV sequences in the polylinker was eliminated by partial digestion of the plasmid PS(1)pΔ2840-2767LUC with *Hind*III, filling in the overhang with the Klenow fragment of *E. coli* DNA polymerase, and religating the plasmid. The plasmids PS(1)pΔ2840-2767(2)HNF1LUC and PS(1)pΔ2840-2767(10)HNF1LUC were constructed by kinasing the same double-stranded oligonucleotide used above to multimerize the oligonucleotide, followed by ligation to the *Hind*III site in PS(1)pΔ2840-2767(H)LUC. PS(1)pΔ2840-2767(2)HNF1LUC contains two HNF1 binding site oligonucleotides, one in each orientation. PS(1)pΔ2840-2767(10)HNF1LUC contains ten HNF1 binding site oligonucleotides, eight in the opposite (-) and two in the same (+) orientation as in the HBV genome. The plasmid PS(1)pΔ2840-2767AHNF1(+)LUC was constructed by inserting a double-stranded oligonucleotide containing the human albumin promoter HNF1 binding site, produced by annealing the oligonucleotides AGCTAGTTAATAATCTACAA and AGCTTTGTAGATTATTAACT, into the unique *Hind*III site in the plasmid PS(1)pΔ2840-2767(H)LUC in the same orientation (+) as the HNF1 binding site in the HBV genome.

The plasmids pHNF1LUC, pHBVTATALUC, and pHNF1TATALUC were constructed by inserting synthetic double-stranded oligonucleotides into sites in the polylinker of p19DLUC. pHNF1LUC was made by inserting a double-stranded oligonucleotide containing the HBV PS(1) promoter HNF1 binding site, produced by annealing the oligonucleotides AGCTAGTTAATCATTACTTC and AGCTGAAGTAATGATTAACT (spanning HBV nucleotides 2719 to 2734), into the unique *Hind*III site of p19DLUC in the same orientation as the HNF1 binding site in the HBV genome. pHBVTATALUC was constructed by inserting a double-stranded oligonucleotide containing the PS(1) promoter TATA box element, produced by

annealing the oligonucleotides CTATATTATATAAGAGAGAAGCT and TCTCTCTTATATAATATAGGTAC (spanning HBV nucleotides 2773 to 2791), into the *Sac*I and *Kpn*I sites of p19DLUC in the same orientation as the TATA box element occurs in the HBV genome. pHNF1TATALUC was made by inserting the double-stranded oligonucleotide containing the PS(1) promoter TATA box element, produced by annealing the oligonucleotides CTATATTATATAAGAGAGAAGCT and TCTCTCTTATATAATATAGGTAC (spanning HBV nucleotides 2773 to 2791), into the *Sac*I and *Kpn*I sites of pHNF1LUC.

A series of constructs was made in which the spacing between the HNF1 site and the TATA box element in the plasmid pHNF1TATALUC was changed by inserting linkers between the two binding sites. The distance between the HNF1 and TATA binding sites in the plasmid pHNF1TATALUC is 39 nucleotides. This distance was altered by digesting the plasmid with *Kpn*I and *Sal*I, and treating with either *T4* DNA polymerase or S1 nuclease to create blunt ends for the insertion of linkers. The plasmids were ligated without linkers or with *Hind*III 8-mers [d(CAAGCTTG)] or 10-mers [d(CCAAGCTTGG)], creating the plasmids pHNF1(10)TATATLUC, pHNF1(12)TATALUC, pHNF1(14)TATALUC, pHNF1(15)TATALUC, pHNF1(18)TATALUC, pHNF1(22)TATALUC, pHNF1(25)TATALUC, pHNF1(26)TATALUC, pHNF1(28)TATALUC, pHNF1(30)TATALUC, pHNF1(35)TATALUC, pHNF1(38)TATALUC, pHNF1(41)TATALUC, pHNF1(45)TATALUC, pHNF1(48)TATALUC, pHNF1(50)TATALUC, pHNF1(58)TATALUC, and pHNF1(66)TATALUC. The number in parentheses indicates the number of nucleotides between the HNF1 and TATA binding sites in the construct. The sequence of each construct was verified by dideoxynucleotide sequencing.

In the plasmid PS(1)pLUC, the regulatory and coding regions for the X gene are intact, and the product of the X gene could be synthesized upon transfection of this plasmid. The plasmids PS(1)pX1LUC and PS(1)pX2LUC are derivatives of PS(1)pLUC designed to remove the coding capacity of the X gene contained in PS(1)pLUC. The plasmid PS(1)pX1LUC was constructed by digestion of PS(1)pLUC with *Nco*I, followed

by digestion with S1 nuclease (Boehringer Mannheim Biochemicals) to remove the four nucleotide CATG overhang (coordinates 1375 to 1378), and religation. This deletion removed the ATG of the translational start site for the X gene. The plasmid PS(1)pX2LUC was constructed by digesting PS(1)pLUC with *Rsr*II, filling in the overhang with the Klenow fragment of *E. coli* DNA polymerase, and ligating a *Sac*I linker [d(CGAGCTCG)]. This generated an 11 bp insertion at coordinate 1577 in the X gene coding region. The regions of the deletion and insertion were confirmed by dideoxynucleotide sequencing.

Plasmids containing the chloramphenicol acetyltransferase (CAT) gene were used in transfection experiments as internal control plasmids. The plasmid pUCCATpA (Linney & Donerly, 1983) contains the promoterless CAT gene in pUC8 and is used for the generation of a negative control extract for the CAT assay. The plasmid pSV2CAT contains the CAT gene under the control of the simian virus 40 (SV40) early promoter (Gorman *et al.*, 1982). Similarly, the plasmid pMTCAT contains the CAT gene under the control of the mouse metallothionein-I promoter. This plasmid was constructed by digesting pUCCATpA with *Hind*III and *Bam*HI, filling in the overhangs with the Klenow fragment of *E. coli* DNA polymerase, ligating an *Xho*I linker [d(CCTCGAGG)], and ligating the 773 bp CAT fragment into the *Sal*I site of pMT. pMT comprises the mouse metallothionein-I promoter sequences from -700 to +64 and the polyadenylation recognition sequence from +930 to +1241 (Glanville *et al.*, 1981; Searle *et al.*, 1984). Blunt ends were created using *T4* DNA polymerase at the *Bgl*II (+64) and *Sac*II (+930) sites and a *Sal*I linker [d(GGTCGACC)] was inserted between the promoter and polyadenylation sequences to create a unique cloning site in pMT (Raney *et al.*, 1990).

The expression vector pMTHNF1 contained the rat HNF1 cDNA (Frain *et al.*, 1989) which was expressed under the control of the mouse metallothionein-I promoter. pMTHNF1 was constructed by ligating the *Bgl*II to *Sac*II fragment containing the HNF1 cDNA into the *Bgl*II and *Sac*II sites at nucleotides +64 and +930 of pUCMT. pUCMT contains the mouse metallothionein-I sequences from -700 to +1241 (Glanville *et al.*, 1981; Searle *et al.*, 1984) cloned into pUC13. The HNF1 cDNA deletion

constructs were produced from the full-length cDNA by using convenient restriction enzyme sites or *Bal*31 nuclease digestion. A termination codon was introduced into the carboxyl-terminal deletions by insertion of an *Nhe*I linker [d(CTAGCTAGCTAG)]. The designations of the truncated HNF1 polypeptides indicate which amino acid residues have been deleted from the 628-amino-acid HNF1 polypeptide. pMTHNF1Δ549-628 was constructed by digestion of pMTHNF1 with *Xba*I and *Stu*I, filling in the overhangs with the Klenow fragment of *E. coli* DNA polymerase, and ligating the *Nhe*I linker. The *Xba*I site is located immediately 3' to the HNF1 cDNA sequences, and the *Stu*I site is located at nucleotide 1685 within the cDNA (Frain *et al.*, 1989). This digestion removed the carboxyl terminus of the cDNA from nucleotide 1685, deleting sequences encoding amino acids 549 to 628. Similarly, pMTHNF1Δ265-628 was constructed by digestion of pMTHNF1 with *Xba*I and *Acc*I, which digests the cDNA at nucleotide 837, followed by a Klenow reaction and ligation of an *Nhe*I linker, resulting in the deletion of amino acids 265 to 628. The plasmids pMTHNF1Δ283-628 and pMTHNF1Δ391-628 were constructed by digestion of the HNF1 cDNA with *Xba*I and *Bal*I, or *Xba*I and *Aat*II, respectively, followed by filling in the overhangs, ligating an *Nhe*I linker, and cloning the deleted fragment into the *Bgl*II and *Sac*II sites of the pUCMT expression vector. The deletion plasmids pMTHNF1Δ283-392 and pMTHNF1Δ283-547 were constructed by digestion of the HNF1 cDNA with *Bal*I and *Aat*II, or *Bal*I and *Stu*I, respectively, followed by filling in the overhangs, self-ligation, selection of the deleted fragment and cloning it into the *Bgl*II and *Sac*II sites of pUCMT. The plasmid pMTHNF1Δ283-392/549-628 was constructed by digestion of pMTHNF1Δ283-392 with *Xba*I and *Stu*I, followed by fill-in and ligation of an *Nhe*I linker. The remaining deletion plasmids, pMTHNF1Δ517-628, pMTHNF1Δ502-628, pMTHNF1Δ482-628, pMTHNF1Δ476-628, pMTHNF1Δ458-628, and pMTHNF1Δ400-628 were generated by digesting pMTHNF1 with *Stu*I, followed by *Bal*31 nuclease digestion according to the manufacturer's instructions. The pool of deletions was then digested with *Xba*I, the overhangs were filled in, and a ligation with *Nhe*I linkers was performed. The deletion breakpoints of these plasmids were confirmed by dideoxynucleotide sequencing (Sanger *et al.*, 1977).

The truncated HNF1 polypeptides were expressed using the mouse metallothionein-I vector pUCMT as described for the full-length HNF1 polypeptide.

VI. B. Cells and Transfections

The human hepatoblastoma cell lines Huh7 (Nakabayashi *et al.*, 1982), HepG2 (Aden *et al.*, 1979) and HepG2.1 (Raney *et al.*, 1990), and the human hepatocellular carcinoma cell lines PLC/PRF/5 (Alexander cells) (Alexander *et al.*, 1976) and Hep3B (Knowles *et al.*, 1980), were grown in RPMI-1640 medium and 10% fetal bovine serum at 37°C in 5% CO_2/air. The human cervical carcinoma cell line HeLa S3 (Puck *et al.*, 1956), and the mouse fibroblast cell line NIH 3T3, were grown in Dulbecco's modified Eagle's medium (DMEM) containing 4.5 mg/ml glucose and 10% calf serum at 37°C in 5% CO_2/air. All of the cell lines were transfected by the calcium phosphate technique, essentially as previously described (Graham & Van der eb, 1973; Sorge *et al.*, 1984). The transfected DNA mixture comprised 15 μg of a LUC plasmid and 1.5 μg of pSV2CAT or pMTCAT, which served as an internal control for transfection efficiency. pSV2CAT and pMTCAT direct the expression of the chloramphenicol acetyltransferase (CAT) gene using the SV40 early promoter and the mouse metallothionein promoter, respectively. When appropriate, the DNA mixture included 1.5 μg of expression vector plasmids. The cells were split the day prior to transfection to achieve a density of approximately 1 to 2 X 10^6 cells per 10 cm^2 plate at the time of transfection. The cells were cultured in fresh DMEM containing 10% calf serum for four hours prior to the addition of the DNA precipitate. The DNA precipitate was prepared by slowly adding 0.5 ml of DNA resuspended in 0.25 M $CaCl_2$ to 0.5 ml of a 2X Hepes buffer (16 g. NaCl, 0.74 g. KCl, 0.25 g. Na_2HPO_4, 10.0 g. Hepes, dH_2O to 1 liter, pH 7.08), gently vortexing after each addition of a few drops of DNA solution. The DNA mixture was incubated 15 minutes before adding it to the plate containing medium and cells. The cells were then incubated at 37°C for 4 to 6 hours before the medium containing the DNA precipitate was aspirated and the cells were incubated at 37°C for two minutes in a solution of 15%

(vol/vol) glycerol in isotonic buffered saline (Parker & Stark, 1979). The cells were then washed with complete medium and incubated in complete medium at 37°C until harvest, 40 to 48 hours later.

VI. C. Preparation of Cell Extracts

Cell extracts were prepared 40 to 48 hours after transfection and assayed for luciferase and CAT activity essentially as previously described (Gorman *et al.*, 1982; De Wet *et al.*, 1987). The cell extracts were prepared by removing the cells from the plate using a cell scraper, pelleting the cells by centrifugation at 1200 revolutions per minute (rpm) (approximately 300 *g*) for 10 minutes using the H-1000B rotor in a Sorvall RT6000 centrifuge, resuspending the pellet in 1 ml phosphate buffered saline (PBS), pelleting the cells by centrifugation for 3 minutes at 10,000 rpm (approximately 8,200 *g*) using a standard 18-place rotor in a Brinkmann microcentrifuge, and resuspending the cells in 100 μl of extract buffer containing 0.1 M KPO_4 (pH 7.8) and 1 mM dithiothreitol (DTT). The cells were lysed by freezing and thawing three times in dry ice and a 37°C water bath. The cells were centrifuged at 14,000 rpm (approximately 16,000 *g*) for 20 to 30 minutes using the standard 18-place rotor in a Brinkmann microcentrifuge and the supernatant (lysate) was collected and stored at -20°C.

The protein concentration of the extracts was determined essentially by the method of Bradford (Bradford, 1976). 20 μl of sample was mixed with 1 ml of Bradford dye reagent and absorbance at 595 nm was measured. A standard curve was generated using bovine serum albumin (BSA) standards. The extract protein concentration was determined relative to the BSA standards.

VI. D. Enzyme Assays

Cell extracts for the γ-glutamyltransferase assays were prepared essentially as described for the luciferase and CAT assays. One plate of untransfected cells

(approximately 5 X 10^6 cells) was lysed in 100 μl of extraction buffer, 0.1 M KPO_4, pH 7.3, by freeze-thawing three times. The cell extract was centrifuged at 20,000 *g* for 30 minutes, and the supernatant removed for enzyme analysis. The γ-glutamyltransferase assays were performed with 10 μl of cell extract using a Kodak Ektachem DT60 analyzer (Eastman Kodak Co., Rochester, N.Y.).

The luciferase assay was performed by adding 20 μl cell extract to 100 μl of assay buffer containing 0.1 M KPO_4 (pH 7.8), 0.015 M $MgSO_4$, 5 mM ATP, and 1 mM DTT in a polystyrene, 12 x 75 mm luminometer cuvette. The substrate, 100 μl of 1 mM luciferin, pH 6.2 [D-luciferin, potassium salt (Analytical Luminescence Laboratory)], was injected by the luminescence photometer and light units were measured for a period of 10 seconds. D-luciferin is a synthetically made compound specific for the firefly luciferase in the following reactions:

$$ATP + Luciferin\ (LH_2) + Luciferase\ (E) \text{ ------> } LH_2\text{-}AMP\text{-}E + PPi$$

$$LH_2\text{-}AMP\text{-}E + O_2 \text{ ----> } AMP + CO_2 + oxyluciferin + Light\ (560\ nm)$$

The chloramphenicol acetyltransferase (CAT) assay was performed by incubating 20 μl cell extract with 66 μl 0.25 M Tris hydrochloride (pH 7.8), 35 μl dH_20, 20 μl 4mM Acetyl CoA, and 4 μl (0.1 μCi) ^{14}C-chloramphenicol [D-threo (dichloroacetyl-1-^{14}C), Amersham] for 30 minutes at 37°C. The reaction results in the transfer of acetyl group(s) from the Acetyl CoA to the substrate chloramphenicol by the CAT enzyme, producing 1-acetyl chloramphenicol, 3-acetyl chloramphenicol and 1,3-diacetyl chloramphenicol (Fig. VI. 3). After the incubation, 1 ml ethyl acetate was added to each tube, and tubes were vortexed briefly after each addition to stop the reaction. Each tube was then vortexed for 30 seconds. The phases were separated by centrifugation at 14,000 rpm (approximately 16,000 *g*) using the standard 18-place rotor in a microcentrifuge for 2 minutes. The top 1 ml organic phase was removed to a clean tube and dried in a Speedvac vacuum centrifuge (Savant) for 45 to 60 minutes. The dried pellet was resuspended in 20 μl ethyl acetate and spotted onto a thin-layer chromatography (TLC) plate (Silica gel, 250 μm). The TLC was run for approximately 45 minutes in a solution of 95% chloroform and 5% methanol after which the plate was

CHLORAMPHENICOL ACETYLTRANSFERASE (CAT)

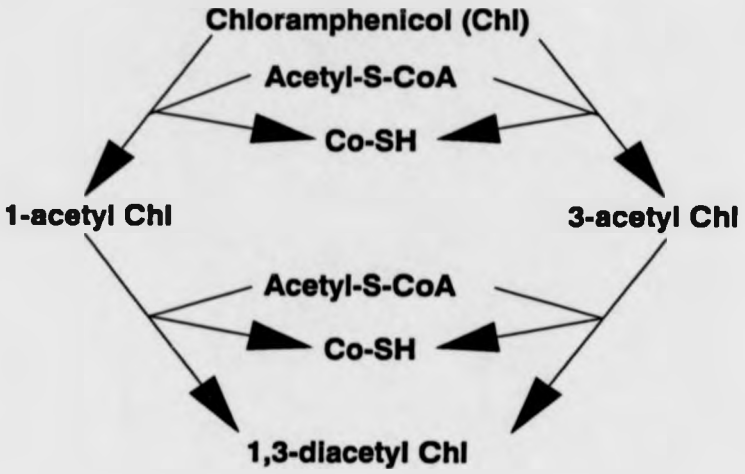


Figure VI.3. Diagrammatic representation of the reaction catalyzed by chloramphenicol acetyltransferase (CAT). The substrate chloramphenicol (Chl) is acetylated at the 1 position, the 3 position, or both, to produce 1-acetyl Chl, 3-acetyl Chl, or 1,3-diacetyl Chl, respectively. The acetyl group is donated by the Acetyl-S-CoA which is reduced to form Co-SH as a byproduct of the reaction.

dried and exposed to X-ray film overnight at room temperature. The assay was quantitated by cutting out the radioactive spots on the TLC plate and counting for 2 minutes in scintillation fluid on a β-scintillation counter.

The luciferase values were normalized to the CAT values for each experiment to correct for different transfection efficiencies of each plate. A Lotus spreadsheet (Fig. VI. 4) was used to calculate the relative luciferase activities shown in the figures. Luciferase values were calculated by subtracting the assay background by subtracting the background from an extract generated from a plate transfected with p19DLUC, and correcting for the dilution factor. The CAT values were calculated by subtracting the background from the TLC plate, and by subtracting the percentage conversion background equivalent to that from an extract generated from a plate of cells transfected with pUCCATpA, a promoterless CAT vector. This calculation is included to correct for ^{14}C-chloramphenicol breakdown products that may cause radioactivity to run into the acetylated portion of the TLC plate. The CAT values were also corrected for the dilution factor. The corrected number of light units produced by the luciferase activity was divided by the corrected percentage of chloramphenicol which was acetylated by the CAT activity in the same cell extract, producing the LUC/CAT ratio. These normalized luciferase activities were reported relative to the activity of the full-length HBV promoter construct, whose value was assigned 1.00 (Raney *et al.*, 1989).

VI. E. Nuclear Extracts

Nuclear extracts were prepared from Huh7 cells and HepG2.1 cells essentially as described previously (Dignam *et al.*, 1983; Raney *et al.*, 1989). All operations were performed at 0 to 4°C. Cells were harvested using cell scrapers and centrifuged for 5 minutes at 2,000 rpm (approximately 833 *g*) using the H-1000B rotor in a Sorvall RT6000 centrifuge. Pelleted cells were washed in 20 ml of 10 mM sodium phosphate (pH 6.8), 0.14 M NaCl, 10 mM $MgCl_2$, by suspension and recentrifugation for 5 minutes at the same speed as used in the first step. The pellet was suspended in five volumes of

Figure VI.4. Lotus spreadsheet used for calculating the relative activities using the luciferase and CAT values from the extracts made after transient transfections. The constructs transfected are listed on the left. The "Luciferase" column contains the number of light units generated by the extract corresponding to the transfected plasmid in the luciferase assay. The "LUC. Backgrd." column contains the value of the luciferase assay background generated by the negative control vector p19DLUC. The "cpm" column contains the counts per minute values from the quantitation of the CAT assay, with the acetylated value on top and the unacetylated value on the bottom. The "CAT Backgrd." column contains the counts per minute value obtained by counting a portion of the thin layer chromatography (TLC) plate that had no sample run on it. This represents the background of the TLC plate. The "cpm-Backgrd." column contains the CAT values after subtraction of the values from the "CAT Bkgd" column. The "Acetyl" column indicates the acetylated (y) or nonacetylated (n) portion of each sample. The "Total cpm" value is the sum of the acetylated and nonacetylated cpm values for each sample. The "Backgrd." column represents each sample's acetylated cpm value equivalent to the percentage of acetylation obtained from the promoterless CAT construct, which represents the background acetylation value that may result from the breakdown of ^{14}C-chloramphenicol. The "Corrected cpm" column contains the cpm values after subtraction of the "Backgrd." value from the acetylated cpm value. The "Total Corr. cpm" column contains the sum of the acetylated and nonacetylated corrected cpm values. The "% Acetyl" column contains the value of the ratio of acetylated cpm to total cpm multiplied by 100. The "% Acetyl. Corr. Fac" column contains the dilution factor of the extract used for the CAT assay. The "Corrected % Acetyl." contains the "% Acetyl" value multiplied by the "% Acetyl. Corr. Fac" value. The "Corrected LUC" column contains the luciferase value after subtraction of the "LUC Backgrd." value. The "LUC Corr. Fac." column contains the dilution factor of the extract used for the luciferase assay. The "LUC light units" column contains the "corrected LUC" value multiplied by the "LUC Corr. Fac." value. The "LUC/CAT" column contains the ratio of the "LUC light units" value to the "Corrected % Acetyl." value. The "Rel. Act." column contains the "LUC/CAT" value divided by the "LUC/CAT" value of the positive control, or full-length promoter construct. The relative activities represent the luciferase values corrected for transfection efficiency (measured as CAT activity) reported relative to the activity of the positive control plasmid in the experiment, usually the full-length promoter construct.

	Luciferase	LUC Backgrd.	cpm	CAT Backgrd.	cpm--Backgrd.	Acetyl	Total cpm	Backgrd.	Corrected cpm	Total Corr. cpm	% Acetyl	% Acetyl Corr. Fac.	Corrected % Acetyl	Corrected LUC	LUC Corr. Fac.	LUC light units	LUC/CAT	Rel. Act.
Hsp68 (extinction			170.50	34.25	136.25	y	122810.15	136.25	0.00	122473.90	0.00	1.00	0.00					
			122506.15	34.25	122473.90	n			122473.90									
Hsp68 SpLUC	349414.00	69.00	29027.00	34.25	28992.75	y	112870.88	125.43	28867.32	112745.45	25.60	1.00	25.60	349345.00	1.00	349345.00	13894.17	1.00
			83912.38	34.25	83878.13	n			83878.13									
Hsp68 SpLUC	349414.00	69.00	29027.00	34.25	28992.75	y	112870.88	125.43	28867.32	112745.45	25.60	1.00	25.60	349345.00	1.00	349345.00	13894.17	1.00
			83912.38	34.25	83878.13	n			83878.13									
Hsp68 p HSVLUC	69.00	69.00	6466.00	34.25	6431.75	y	110119.75	122.37	6309.38	109996.38	5.74	1.00	5.74	0.00	1.00	0.00	0.00	0.00
			103721.25	34.25	103687.00	n			103687.00									
Hsp68 XpLUC	29883.00	69.00	11870.50	34.25	11836.25	y	120401.99	133.80	11702.45	120268.19	9.73	1.00	9.73	29956.00	1.00	29956.00	2708.39	0.02
			106599.99	34.25	106565.74	n			106565.74									
Hsp68 CpLUC	40406.00	69.00	12925.50	34.25	12891.25	y	118474.77	131.65	12759.60	118343.12	10.78	1.00	10.78	40397.00	1.00	40397.00	4575.92	0.34
			105617.77	34.25	105583.52	n			105583.52									
Hsp68 PE(YpLUC	6900.00	69.00	8174.50	34.25	8140.25	y	124690.29	138.56	8001.69	124551.73	6.42	1.00	6.42	6934.00	1.00	6934.00	1080.29	0.08
			116584.29	34.25	116550.04	n			116550.04									

hypotonic buffer (10 mM Tris hydrochloride [pH 8.0], 10 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM dithiothreitol [DTT], 0.5 mM phenylmethylsulfonyl fluoride [PMSF]), and was allowed to incubate on ice for 10 minutes. The cells were pelleted by centrifugation for 5 minutes at 2,000 rpm (approximately 833 *g*) using the H-1000B rotor in a Sorvall RT6000 centrifuge and resuspended in two volumes of the hypotonic buffer. The cells were then lysed by 10 to 15 strokes of a Dounce homogenizer (B-type pestle) and checked on a microscope for more than 90% lysis. The lysate was centrifuged for 10 minutes at 3,000 rpm in a Sorvall SS-34 rotor, and the supernatant was removed, leaving a loose nuclear pellet. The nuclear pellet was recentrifuged for 20 minutes at 15,000 rpm in a Sorvall SS-34 rotor, and the supernatant was decanted. The nuclear pellet was resuspended in 2.5 volumes of nuclear extraction buffer (20 mM Tris hydrochloride [pH 8.0], 0.42 M NaCl, 1.5 mM $MgCl_2$, 0.2 mM N,N,N',N'-ethylenediaminetetraacetic acid [EDTA], 25% [vol/vol] glycerol, 0.5 mM DTT, 0.5 mM PMSF), and the nuclei were lysed by 10 strokes of the Dounce homogenizer. The lysate was stirred gently for 30 minutes on ice, then centrifuged for 30 minutes at 15,000 rpm in a Sorvall SS-34 rotor. The supernatant was dialyzed for five hours against 50 volumes of dialysis buffer (20 mM Tris hydrochloride [pH 8.0], 100 mM KCl, 0.2 mM EDTA, 20% [vol/vol] glycerol, 0.5 mM DTT, 0.5 mM PMSF) with one change of buffer after 2.5 hours. The nuclear extract was clarified after dialysis by centrifugation for 5 minutes at 14,000 rpm (approximately 16,000 *g*) using the standard 18-place rotor in a Brinkmann microcentrifuge. The extract was aliquoted, frozen in liquid N_2, and stored at -80°C.

A modification of this procedure was used to prepare nuclear extracts from single plates of transfected cells. The cells were harvested 40 to 48 hours after transfection and collected as above. The cells were washed in 1 ml of sodium phosphate (pH 6.8), 0.14 M NaCl, 1.5mM $MgCl_2$, by suspension and centrifugation in a Brinkmann microcentrifuge at 14,000 rpm (approximately 16,000 *g*) using the standard 18-place rotor for 2 minutes. The cell pellet was resuspended in 200 µl of the hypotonic buffer and allowed to stand on ice for 10 minutes. The cells were pelleted by centrifugation in a Brinkmann microcentrifuge at 2,000 rpm (approximately 325 *g*) using the standard 18-

place rotor, resuspended in 100 μl of hypotonic buffer and lysed by 20 to 25 strokes of a Kontes pellet pestle in a microcentrifuge tube. Lysis was more than 90%. The lysate was centrifuged for 10 minutes in the Brinkmann microcentrifuge at 4,000 rpm (approximately 1,300 *g*) using the standard 18-place rotor, the supernatant was removed and the nuclear pellet recentrifuged for 20 minutes at 25,000 rpm (approximately 25,800 *g*) using the TLA-100.3 rotor in a Beckman TL-100 ultracentrifuge. The supernatant was decanted and the nuclear pellet resuspended in 100 μl of nuclear extraction buffer. The nuclear pellet was homogenized by 10 to 15 strokes of the Kontes pellet pestle and the homogenate was shaken gently for 30 minutes then centrifuged for 30 minutes at 25,000 rpm (approximately 25,800 *g*) using the TLA-100.3 rotor in the Beckman TL-100 ultracentrifuge. The nuclear extract (approximately 100 μl) was centrifuged in a pre-rinsed Centricon 10 microconcentrating chamber (Amicon, molecular weight cut-off 10,000) with 2 ml of dialysis buffer for 2 hours at 6,500 rpm in a Sorvall SS-34 rotor. The extract was mixed with 1 ml of dialysis buffer and centrifuged in the Centricon 10 for 3 hours at 6,500 rpm. The sample was collected by centrifugation for 2 minutes at 3,000 rpm in a Sorvall SS-34 rotor. The extract was transferred to a microcentrifuge tube and clarified by centrifugation for 5 minutes at 14,000 rpm (approximately 16,000 *g*) using the standard 18-place rotor in a Brinkmann microcentrifuge. The nuclear extract was aliquoted and frozen in liquid N_2 and stored at -80°C.

VI. F. Gel Mobility Shift Assay

Gel mobility shift analysis was performed essentially as described (Ausubel *et al.*, 1987) using nuclear extracts prepared as described above from untransfected Huh7, HepG2.1 and HeLa cells, and from HepG2.1 cells transfected with 30 μg of the expression vectors encoding the HNF1 polypeptide or various truncated HNF1 polypeptides. The probe, a ^{32}P-labelled double-stranded large surface antigen promoter HNF1 binding site oligonucleotide, was prepared by annealing the oligonucleotides AGCTAGTTAATCATTACTTC and AGCTGAAGTAATGATTAACT to generate the

double-stranded oligonucleotide with AGCT overhangs. 100 ng of this double-stranded oligonucleotide was radiolabelled with 50 μCi α^{32}P-dATP using the Klenow fragment of *E. coli* DNA polymerase. Two, six or 12 μg nuclear extract (Fig. VII. 9, 16) were incubated with 1 μg poly[d(I-C)] in the presence of 25 mM Hepes (pH 7.9), 40 mM KCl, 7.5% (vol/vol) glycerol, 0.13 M EDTA, and 5 mM $MgCl_2$ in a total volume of 20 μl for 15 minutes on ice prior to the addition of 10,000 cpm (approximately 0.05 ng) of ^{32}P-labelled double-stranded HNF1 binding site oligonucleotide. The incubation continued for 15 minutes on ice, then 15 minutes at room temperature before the addition of 2 μl of loading dye [0.25% (wt/vol) bromophenol blue, 0.25% (wt/vol) xylene cyanol, 25% (wt/vol) Ficoll]. The gel mobility shift competition assays were performed as described for the uncompeted reactions with the addition of 1 ng, 10 ng, or 100 ng (20; 200; or 2,000 fold molar excess) unlabelled double-stranded HNF1 binding site oligonucleotide incubated with the 1 μg poly[d(I-C)] 15 minutes prior to the addition of the labelled oligonucleotide. The entire reaction was electrophoresed on a 4% polyacrylamide gel containing 2.5% (vol/vol) glycerol in 0.5 X TBE (0.045 M Tris-borate, 0.045 M boric acid, 0.001 M EDTA) for approximately 3 hours at 75 volts. The gel was fixed for 15 minutes in 10% (vol/vol) glacial acetic acid and 5% (vol/vol) glycerol, dried and exposed to X-ray film.

VI. G. DNase I Footprinting Assay

The DNase I footprinting assay was performed essentially as described previously (Raney *et al.*, 1989; Briggs *et al.*, 1986; Raney *et al.*, 1992). A DNA fragment containing the HBV PS(1) promoter *Hind*III fragment was radiolabelled using 150 μCi γ^{32}P-ATP and *T4* polynucleotide kinase. A 423-bp *Bgl*II to *Hind*III fragment was gel isolated and used as the probe in the footprinting reactions. The reactions contained 1 to 5 ng of end-labelled DNA fragment in a 50-μl reaction mixture containing 25 mM Tris hydrochloride (pH 7.9), 6.25 mM $MgCl_2$, 0.5 mM EDTA, 0.5 mM DTT, 50 mM KCl, 10% (vol/vol) glycerol, and 4 footprinting units (fpu) of purified TATA binding protein

(TBP) (TFIID, Promega), or 1 μg poly[d(I-C)] and 50 μg nuclear extract to be assayed. Binding was carried out for 15 minutes at 0°C and then for 2 minutes at room temperature, after which 50 μl of 5 mM $CaCl_2$, 10 mM $MgCl_2$, containing DNase I, was added at room temperature for 1 minute, and the reaction was stopped by the addition of 100 μl of 1% (wt/vol) sodium dodecyl sulfate (SDS), 20 mM EDTA, 200 mM NaCl, containing 250 μg of tRNA per ml. The quantity of DNase I used was dependent upon the amount of purified protein or nuclear extract in the assay. When no purified protein or extract was used, 0.014 units DNase I (a unit is defined by the manufacturer, Boehringer Mannheim Biochemicals, as the amount of enzyme activity that results in an increase of absorbance of 0.001/minute) was added, when 4 units of TBP was used, 0.288 units DNase I was added, when 50 μg Huh7 nuclear extracts were used, 0.46 units DNase was added, and when 50 μg HepG2.1 or HeLa nuclear extracts were used, 2.3 units DNase I was added. The mixture was then phenol extracted, precipitated with ethanol and analyzed by 6% urea-acrylamide sequencing gel electrophoresis and autoradiography. The DNase I footprinting assay was performed by Alan McLachlan.

VII. RESULTS

VII. A. Control Experiments

The experiments determining the regulatory elements governing transcription of the HBV large surface antigen gene were performed using transient transfections of plasmid DNA into mammalian cell lines in tissue culture, followed by whole-cell lysis to extract proteins, and determination of reporter enzyme activity as a measure of the transcriptional activity of the template transfected. The calcium phosphate method of transfection was performed essentially as described (Graham & Van der eb, 1973) (see section VI. B). In addition to the test plasmid, a second plasmid was always transfected as an internal control for the transfection efficiency of each plate of cells. In this analysis, a promoterless vector containing the ORF of the firefly luciferase gene (p19DLUC, Fig. VI. 1) was used as the reporter molecule for the promoter being analyzed. A vector containing the ORF of the bacterial enzyme chloramphenicol acetyltransferase (CAT) was used as the internal control plasmid (see section VI. A). The expression of the reporter enzyme CAT was under the control of the simian virus 40 (SV40) early promoter (pSV2CAT) (Gorman *et al.*, 1982) or the mouse metallothionein-I promoter (pMTCAT, see section VI. A). The enzyme assays have been described (section VI. D).

Control experiments were performed to examine the limitations of the assays being used. The protein concentration of the cell extract was determined by the method of Bradford (Bradford, 1976). The typical protein concentration of an extract prepared from one 10-cm^2 dish of confluent cells was approximately 2 mg/ml (+/- 1 mg/ml). As the protein concentration did not vary greatly between plates, and the internal control plasmid was transfected to normalize for transfection efficiencies and any differences in the handling of each extract, protein concentration differences were not considered to be a large factor in the interpretation of the data. However, to demonstrate that the enzyme assays were not influenced by the protein concentration of the extracts, serial

dilutions of the extracts were measured in both the luciferase and CAT assays to demonstrate linearity in the assay over a range of protein concentrations (Fig. VII. 1 and 2A). The results shown in Fig. VII. 1 demonstrated that the luciferase assay is linear at least over a range of light units from 3 X 10^6 down to 1.5 X 10^2, and over a range of protein concentrations from 1.7 mg/ml (20 μl extract volume) to 3.3 μg/ml (0.039 μl extract volume). An extract with high relative activity (XpLUC in Huh7 cells) was used to demonstrate linearity of light unit output. These data indicated that the small variation observed in the protein concentrations of the cellular extracts should not substantially affect the luciferase assay. The manufacturer of the luminescence photometer (Analytical Luminescence Laboratory) claims linearity of output up to 1 X 10^7 light units. All of the data generated for the large surface antigen promoter analysis were based on luciferase values of no greater than 1 X 10^6 light units and no less than twice the number of light units generated by the negative control plasmid, p19DLUC, which was typically 150 to 200 light units.

The results of the CAT assay using serial dilutions of the extract (Fig. VII. 2A) demonstrated that the CAT assay is approximately linear over a range of extract protein concentrations from 1.7 mg/ml to 53 μg/ml, indicating that the variation in protein concentration observed between extracts in these analyses should not affect the CAT assay. Serial dilutions of an extract having a protein concentration within the linear range and with high levels of CAT activity were also measured to determine the range of linearity in terms of percent acetylation by the extract. These results (Fig. VII. 2B) demonstrated that the CAT assay was approximately linear up to 80% acetylation. The CAT assay appeared to reach saturation above 80% acetylation. The insert in Fig. VII. 2B shows the linear fit of the data up to 81% acetylation. All of the data generated for the PS(1) promoter analysis was based on CAT activities between 1% and 80% acetylation. The possibility existed that the promoters being used to direct the expression of the CAT gene could have different transcriptional activities in different cell lines used in these analyses. If the activity in one cell line were very low, the time of the CAT reaction may need to be extended to observe CAT activity. To determine

dilutions of the extracts were measured in both the luciferase and CAT assays to demonstrate linearity in the assay over a range of protein concentrations (Fig. VII. 1 and 2A). The results shown in Fig. VII. 1 demonstrated that the luciferase assay is linear at least over a range of light units from 3 X 10^6 down to 1.5 X 10^2, and over a range of protein concentrations from 1.7 mg/ml (20 μl extract volume) to 3.3 μg/ml (0.039 μl extract volume). An extract with high relative activity (XpLUC in Huh7 cells) was used to demonstrate linearity of light unit output. These data indicated that the small variation observed in the protein concentrations of the cellular extracts should not substantially affect the luciferase assay. The manufacturer of the luminescence photometer (Analytical Luminescence Laboratory) claims linearity of output up to 1 X 10^7 light units. All of the data generated for the large surface antigen promoter analysis were based on luciferase values of no greater than 1 X 10^6 light units and no less than twice the number of light units generated by the negative control plasmid, p19DLUC, which was typically 150 to 200 light units.

The results of the CAT assay using serial dilutions of the extract (Fig. VII. 2A) demonstrated that the CAT assay is approximately linear over a range of extract protein concentrations from 1.7 mg/ml to 53 μg/ml, indicating that the variation in protein concentration observed between extracts in these analyses should not affect the CAT assay. Serial dilutions of an extract having a protein concentration within the linear range and with high levels of CAT activity were also measured to determine the range of linearity in terms of percent acetylation by the extract. These results (Fig. VII. 2B) demonstrated that the CAT assay was approximately linear up to 80% acetylation. The CAT assay appeared to reach saturation above 80% acetylation. The insert in Fig. VII. 2B shows the linear fit of the data up to 81% acetylation. All of the data generated for the PS(1) promoter analysis was based on CAT activities between 1% and 80% acetylation. The possibility existed that the promoters being used to direct the expression of the CAT gene could have different transcriptional activities in different cell lines used in these analyses. If the activity in one cell line were very low, the time of the CAT reaction may need to be extended to observe CAT activity. To determine

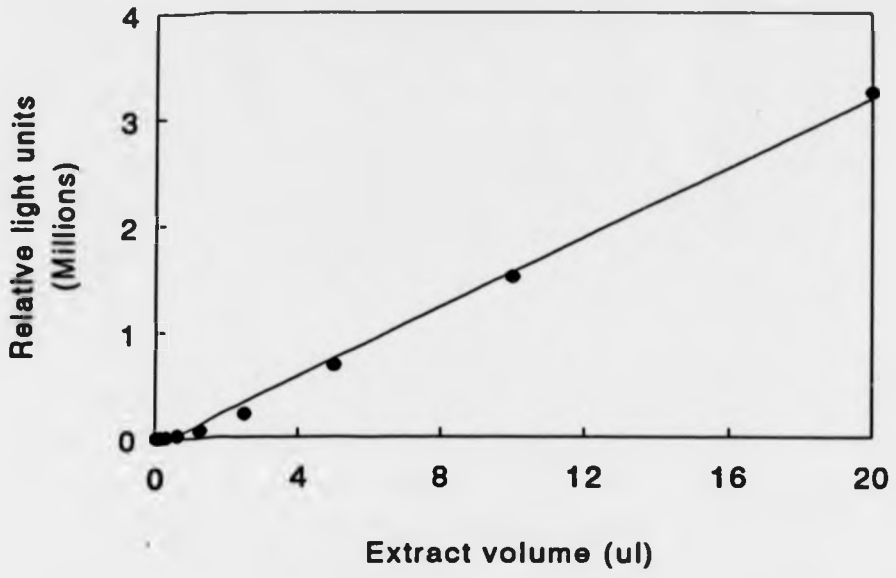
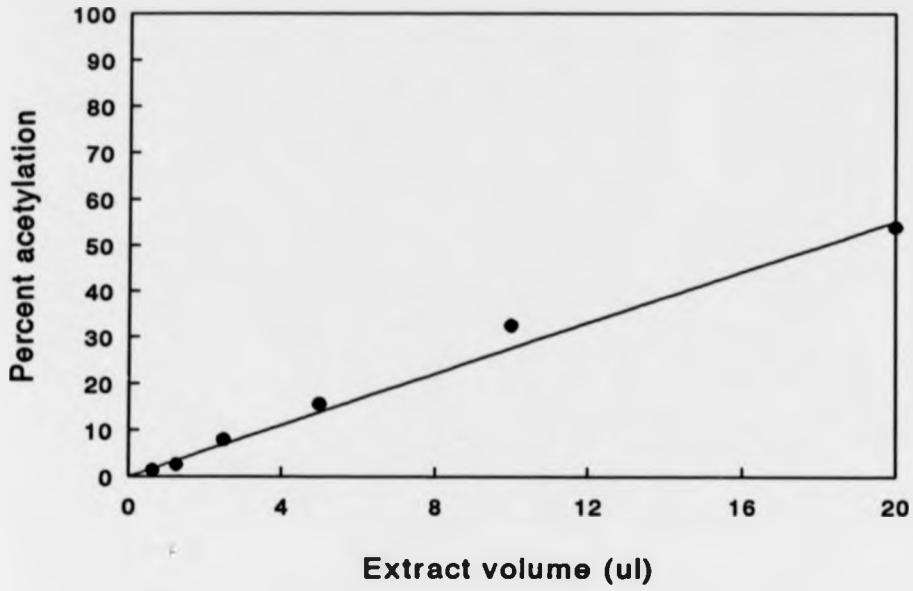


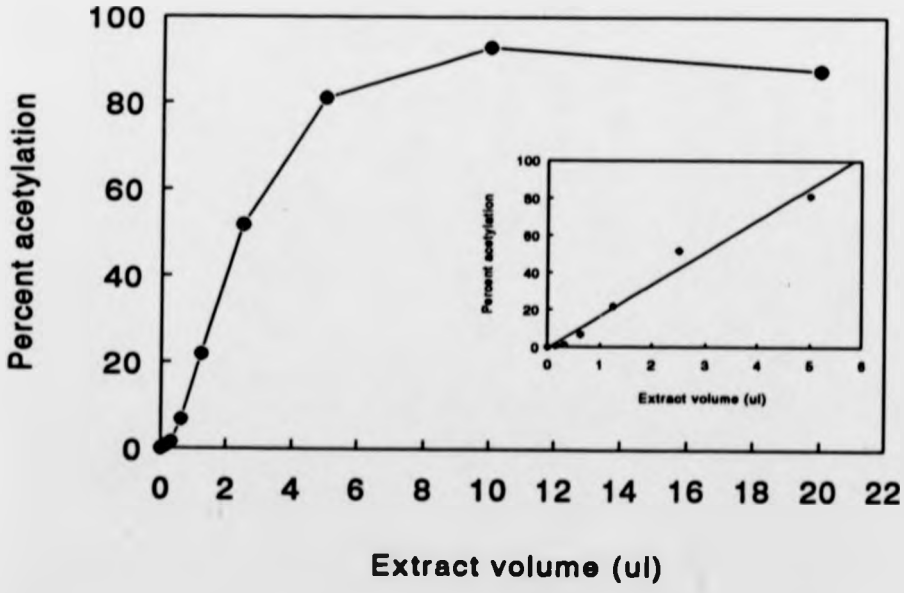
Figure VII.1. Linearity of luminometer output over a serial dilution of luciferase extract. Relative light units are the values obtained after subtraction of the luciferase assay background from the relative light unit output of the luminometer. The extract volume given in μl represents the amount of extract used in the assay.

Figure VII.2. (A) and (B) Linearity of chloramphenicol acetyltransferase (CAT) assay over a serial dilution of CAT extract. Percent acetylation represents the value of the ratio of the acetylated counts per minute (cpm) to the total cpm multiplied by 100. The extract volume (μl) represents the amount of extract used in the assay. (C) Linearity of CAT assay over time. The percent acetylation is calculated as described in section (A) and is plotted relative to the length of time (in hours) the CAT reaction proceeded. (D) Diagrammatic representation of the thin layer chromatography (TLC) portion of the CAT assay. The percent acetylation is calculated as described in section (A). The Origin line represents the position on the TLC plate where the samples were applied. The direction of migration is indicated by an arrow. The CAT reaction products are labelled. Chl, chloramphenicol; 1-acetyl Chl, chloramphenicol acetylated once at the 1 position; 3-acetyl Chl, chloramphenicol acetylated once at the 3 position; 1,2-diacetyl Chl, chloramphenicol acetylated twice at the 1 and 3 positions.

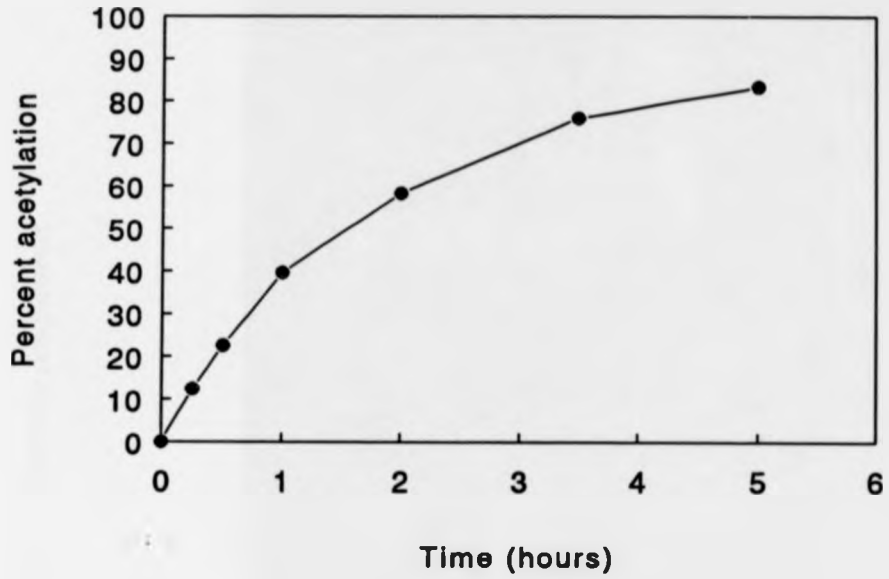
A.


Percent acetylation
100
90
80
70
60
50
40
30
20
10
0
0
4
8
12
16
20
Extract volume (ul)

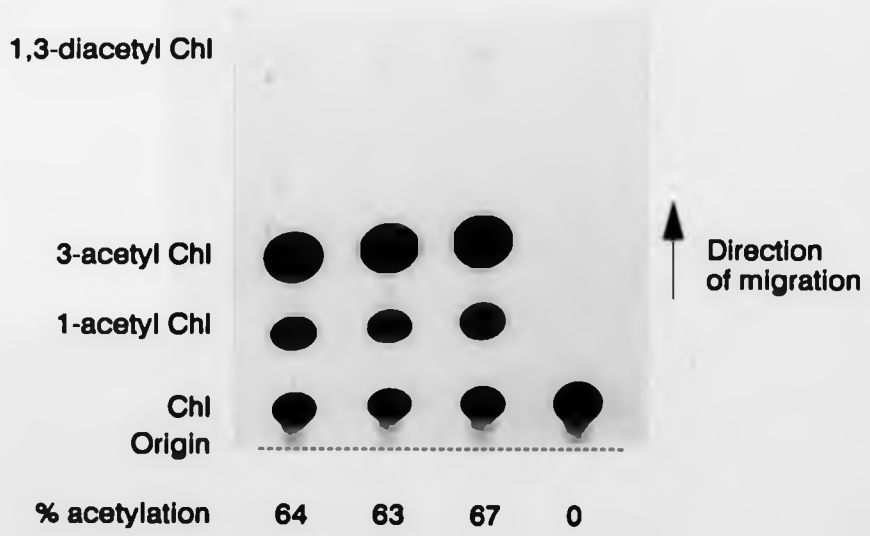
B.


Percent acetylation
100
80
60
40
20
0
0
2
4
6
8
10
12
14
16
18
20
22
Extract volume (ul)
Percent acetylation
Extract volume (ul)

C.



D. CAT TLC ASSAY



the linearity of the CAT assay as a function of time, several time points during a CAT assay were measured. These data (Fig. VII. 2C) indicated that the CAT assay was approximately linear for at least one hour under the standard conditions of the assay (see section VI. D) and curvilinear up to five hours. Although the CAT assay was not completely linear up to five hours, the relative luciferase/CAT ratios between samples remained constant as long as every sample was tested for CAT activity for the same length of time, from 15 minutes to 5 hours. The length of time the CAT assays were run for the large surface antigen promoter experiments was 30 minutes. An autoradiograph of a typical CAT assay is shown in Fig. VII. 2D.

Another possible variable of the transient transfection system is the actual amount of DNA that is taken up by the population of cells being transfected. The preparation of DNA could affect the formation of the DNA-$CaPO_4$ precipitate, and the condition of the cells could affect the ability of the DNA precipitate to be taken up by the cells. The inclusion of an internal control plasmid in all of the transfections should correct for such differences in transfection efficiency. To determine whether the corrected luciferase values (luciferase/CAT ratio) were affected by differences in the amount of DNA transfected into the cells, an experiment was performed in which increasing amounts of DNA were transfected into the cells. The relative amount of luciferase and CAT plasmids transfected into each dish of cells remained constant, at a 10:1 ratio, which was the ratio routinely used for the promoter experiments. The results shown in Table VII. 1 demonstrated that the luciferase and CAT values increased as the amount of DNA transfected was increased, but the luciferase/CAT ratio remained approximately the same, in a range from 0.87 to 1.10 around a relative activity of 1.00 for the standard amount of DNA transfected. It appears that the corrected luciferase values (luciferase/CAT ratio) remains constant regardless of the absolute amount of DNA transfected into a dish of cells, indicating that inclusion of a control plasmid in each transfection will correct for differences of amounts of DNA that reach the cell.

The majority of experimental variation can be corrected for by including an internal control in each experiment and by performing the assays within the their linear

TABLE VII.1 Effect of DNA concentration on the relative luciferase activity[a].

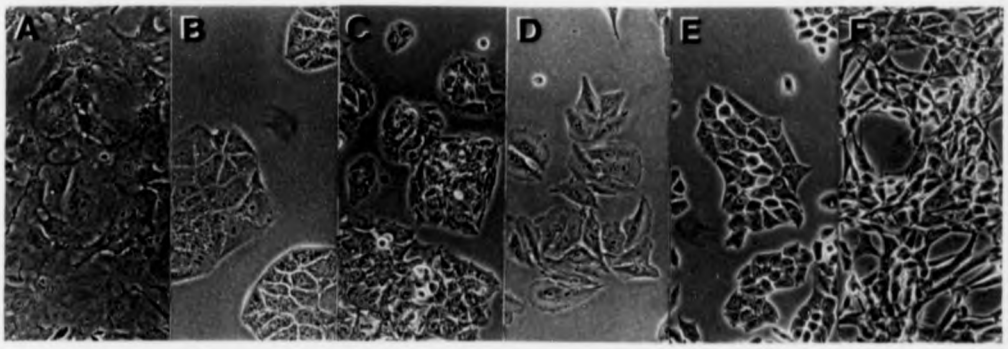
Amt. of DNA Transfected		LUC Light Units	CAT % Acetylation	LUC/CAT Ratio	Relative Activity
LUC	CAT				
3.0	0.30	342,133	2.81	121,755	1.10
7.5	0.75	1,597,839	15.85	100,810	0.91
10.0	1.0	2,202,058	20.96	105,060	0.95
15.0	1.5	2,997,565	27.11	110,570	1.00
20.0	2.0	2,632,937	27.36	96,233	0.87
30.0	3.0	3,080,718	28.50	108,095	0.98

[a]Increasing amounts of DNA were transfected into Huh7 cells and the relative activities were calculated as the ratio of the luciferase light units value to the CAT % acetylation value. The ratio of LUC DNA to CAT DNA transfected was always 10:1. The activities were reported relative to the activity of the transfection containing 15 μg of LUC DNA and 1.5 μg of CAT, which are the standard amounts used for the experiments.

range. However, a degree of variability still exists in the results of transfection experiments using different plasmid preparations of the same construct. This variation, which could be due to the purity and quality of the preparation, is generally within two-fold, but is sometimes greater. For this reason, most of the transfection experiments are repeated with different preparations of the plasmids to obtain an average value. Additionally, the transfection experiments defining critical breakpoints in the determination of important regions for promoter activity were confirmed by alternative experiments, such as internal mutations and deletions to complement a deletion series, the use of synthetic oligonucleotide minimal promoter constructs, and gel mobility shift analysis to provide physical data supporting the functional data.

VII. B. Characterization of Cell Lines

The functional analysis of the transcriptional regulatory elements responsible for HBV large surface antigen gene expression was performed by transient transfection assays using several mammalian cell lines. The human hepatoblastoma cell lines Huh7, HepG2, and HepG2.1, the human hepatocellular carcinoma cell lines PLC/PRF/5 (Alexander) and Hep3B, the human cervical carcinoma cell line HeLa S3, and the murine fibroblast cell line NIH 3T3, were used. The appearances of these cells in culture were examined (Fig. VII. 3). The differentiated human hepatoma cell lines HepG2, PLC/PRF/5, and Hep3B displayed a regular hepatocyte-like epithelial morphology in culture. The cells proliferated in distinct patches rather than spreading evenly over the surface of the tissue culture dish. In contrast, the dedifferentiated hepatoma cell line HepG2.1, which arose spontaneously in culture from the differentiated hepatoma cell line HepG2, displayed a flat, irregular shape and did not form large patches of cells, but spread over the plate in a more fibroblast-like manner. The morphological characteristics of the HepG2.1 cell line became apparent over a period of several months of culturing the HepG2 cells. The majority of the population of cells lost the morphological characteristics of a differentiated hepatoma cell line (Raney *et al.*, 1990).



Cell line	Hep3B	PLC/PRF/5	HepG2	HepG2.1	HeLa	3T3
GGT (mIU /mg)	26.7	23.2	25.4	2.6	1.0	0.3

Figure VII.3. **Phase contrast photomicrographs (200X magnification) of living cultures of Hep3B (A), PLC/PRF/5 (B), HepG2 (C), HepG2.1 (D), HeLa S3 (E), and NIH 3T3 (F) cells. The γ-glutamyltransferase (GGT) specific activities (in milli-international units [mIU] per milligram of protein) in extracts from these cells are indicated.**

The appearances of the PLC/PRF/5 and Hep3B cells did not change under the same culture conditions. The Huh7 cell line, which was brought into the laboratory after the initial analysis, also displayed a regular, hepatocyte-like morphology. The Huh7 cells appeared to be stable in culture and were therefore used for much of the subsequent analysis of the large surface antigen promoter. Like the HepG2 cell line, the Huh7 cell line has been shown to support HBV replication (Sureau *et al.*, 1986; Tsurimoto *et al.*, 1987; Sells *et al.*, 1987; Chang *et al.*, 1987; Yaginuma *et al.*, 1987), suggesting that these two cell lines may represent the tissue culture system most closely resembling the natural environment of the virus. The HeLa S3 cells displayed the typical epithelial cell type morphology, and the NIH 3T3 cells displayed the characteristic fibroblast cell type morphology.

In hepatoma cell lines the levels of various liver enzyme activities are measured to determine the differentiation state of the cell lines. The levels of γ-glutamyltransferase activity in the cell lines used for transfections were measured as an indication of the differentiation state of the various hepatoma cell lines (Richards *et al.*, 1982) (see section VI. D). The γ-glutamyltransferase activity was high in the differentiated hepatoma cell lines, very low in the non-hepatoma cell lines, and somewhat higher in the dedifferentiated hepatoma cell line HepG2.1 (Fig. VII. 3). This result demonstrated a biochemical difference in addition to the morphological difference between the differentiated hepatoma, dedifferentiated hepatoma, and non-hepatoma cell lines.

VII. C. Relative Activities of the Four HBV Promoters

The predominant RNA transcripts observed during HBV infection are the 2.1 kb and the 3.5 kb mRNAs, controlled by the major surface antigen (Sp) and core, or nucleocapsid, promoters (Cp), respectively (Cattaneo *et al.*, 1984; Yokosuka *et al.*, 1986; Su *et al.*, 1989b). The 2.4 kb mRNA, controlled by the large surface antigen promoter [PS(1) promoter], is observed at much lower levels than the 2.1 and 3.5 kb mRNAs. The

0.7 kb mRNA encoding the X gene product has not been observed during HBV infection (Cattaneo *et al.*, 1984; Yokosuka *et al.*, 1986; Su *et al.*, 1989b) but has been detected in WHV infection (Kaneko & Miller, 1988) and in transgenic mice (Araki *et al.*, 1989) and cell culture (Zelent *et al.*, 1987a), and a promoter region upstream of the gene has been identified (Siddiqui *et al.*, 1987; Treinin & Laub, 1987; Zhang *et al.*, 1992). In order to test the relative strengths of the four viral promoters in the context of the complete viral genome, the promoter-LUC constructs containing the complete HBV genome (Fig. VII. 4) were transfected into several cell lines, and their transcriptional activities were examined (Table VII. 2). The *ayw* subtype of HBV was used for all of the analyses because it has been shown that this cloned DNA can produce infectious viral particles (Will *et al.*, 1982; Will *et al.*, 1985), indicating that the sequence of this subtype contains all of the necessary elements for the function of the virus. The SpLUC, XpLUC, CpLUC, and PS(1)pLUC constructs were transfected into the differentiated hepatoma cell lines Huh7, Hep3B, PLC/PRF/5, and HepG2, the dedifferentiated hepatoma cell line HepG2.1, and the nonhepatoma cell lines HeLa S3 and NIH 3T3. The relative activities were determined by calculating the ratios of the normalized luciferase activities of XpLUC, CpLUC, and PS(1)pLUC to the normalized luciferase activity of SpLUC in each cell line (see section VI. D and Fig. VI. 4 as an example of the calculations). The activities were reported relative to the activity of the S promoter in each cell line because the activity of the internal control promoter may vary in the different cell lines. The internal control transfected in these experiments was pSV2CAT.

The relative levels of activity of the four HBV promoters were compared in the seven cell lines (Table VII. 2). All four HBV promoters exhibited some level of transcriptional activity in each cell line. The level of activity of the X promoter was similar to that of the S promoter in all of the cell lines except Huh7, in which the level of X promoter activity was six-fold that of the S promoter. The Huh7 cell line may contain a transcription factor(s) that enhances the X promoter activity which is not present in the other cell lines. The relative level of activity of the X promoter in these cell lines does not correlate with the low or undetectable level of X gene RNA species

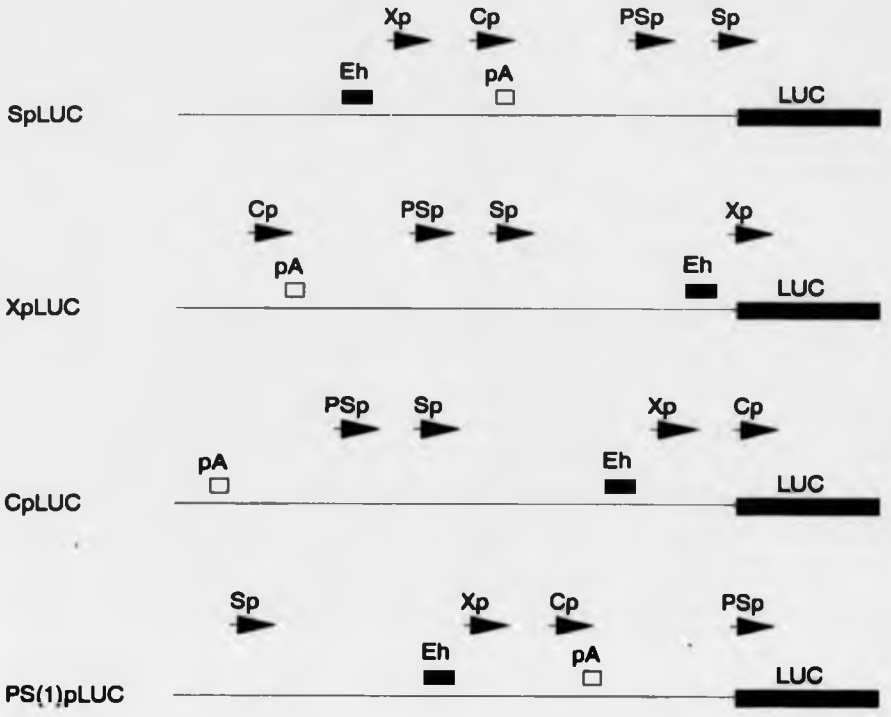


Figure VII.4. The luciferase reporter gene constructs used to determine the relative transcriptional activities from the four HBV promoters. The SpLUC, XpLUC, CpLUC and PS(1)pLUC constructs contain the complete HBV genome inserted into p19DLUC immediately 5' to the luciferase reporter gene such that expression of the luciferase reporter gene is controlled by the major surface antigen, X gene, nucleocapsid (core) and large surface antigen [PS(1)] promoters, respectively. Arrows indicate the position and direction of transcription from the HBV X gene (Xp), core (Cp), PS(1) (PSp) and major surface antigen (Sp) promoters. Boxes indicate the positions of the HBV enhancer I sequence (Eh), HBV polyadenylation recognition sequence (pA) and the luciferase (LUC) open reading frame. The horizontal lines indicate the HBV sequences present in the various luciferase constructs.

TABLE VII. 2. Relative activities of the HBV promoters[a].

HBV Promoter	Cell lines						
	Huh7	Hep3B	PLC/ PRF/5	HepG2	HepG2.1	HeLa	NIH 3T3
SpLUC	1.00	1.00	1.00	1.00	1.00	1.00	1.00
XpLUC	6.21	2.02	1.33	0.66	0.65	2.44	0.88
CpLUC	5.17	0.34	1.35	0.90	0.15	0.20	0.068
PS(1)pLUC	0.14	0.078	0.016	0.046	0.0019	0.0028	0.0009

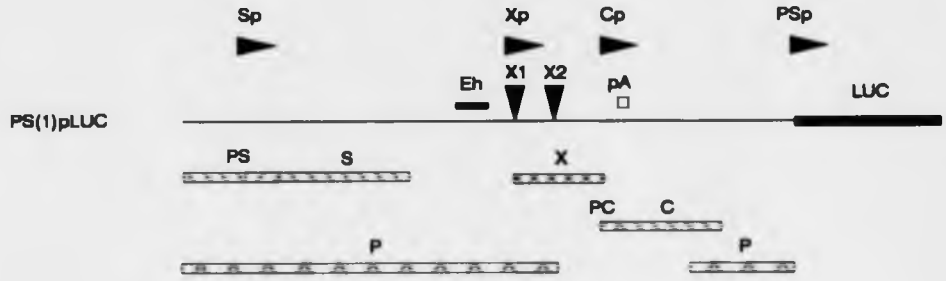
[a] The activities of the HBV promoters are reported relative to the activity of the major surface antigen promoter in each cell line. The internal control used to correct for transfection efficiencies was pSV2CAT. The luciferase activities were normalized by dividing the corrected number of light units produced by the luciferase activity by the corrected percentage of chloramphenicol which was acetylated by the CAT activity in the same cell extract. The relative activities were determined by calculating the ratios of the normalized luciferase activities of XpLUC, CpLUC, and PS(1)pLUC to the normalized luciferase activity of SpLUC.

during HBV infection. Some possible explanations for this include a difference in the stability of the luciferase RNA compared with the stability of the X RNA, or differences between the levels of transcription factors regulating the X promoter in the human liver versus the cell lines analyzed in this study. The relative level of activity of the C promoter was similar to the S promoter in the differentiated hepatoma cell lines, with the exception of Huh7, in which the level of C promoter activity was five-fold higher than that of the S promoter. The Huh7 cell line may contain a transcription factor(s) which positively regulates the C promoter and is absent from the other cell lines. The relative level of activity of the C promoter was five- to 75-fold lower in the dedifferentiated hepatoma and nonhepatoma cell lines, suggesting some preferential transcriptional activity from the C promoter in differentiated hepatoma cell lines. The PS(1) promoter exhibited the lowest relative activity of all the promoters in every cell line, ranging from 0.09% to 14% of the level of activity of the S promoter, depending on the cell line. This result is consistent with the relative amounts of major surface, core, and large surface antigen RNAs observed during HBV infection. The PS(1) promoter displayed higher relative activity in the differentiated hepatoma cell lines compared with the dedifferentiated hepatoma and nonhepatoma cell lines. The relative level of activity of the PS(1) promoter was approximately six to 150 times higher in the differentiated hepatoma cell lines than in the other cell lines. The observation that the PS(1) promoter appears to display preferential transcriptional activity in differentiated hepatoma cell lines suggests a possible role for the transcriptional regulation of the large surface antigen gene in the hepatotropism of the virus. The relative weakness of the PS(1) promoter in all of the cell lines may indicate that synthesis of the large surface antigen, an essential component of the virion, could represent a limiting step in the assembly of the Dane particle in cell culture.

VII. D. Influence of the X Gene on the Large Surface Antigen Promoter

It has been reported that the product of the X gene can transactivate several heterologous promoters, including the SV40 enhancer and early promoter (Spandau &

Lee, 1988; Zahm *et al.*, 1988; Twu & Robinson, 1989; Siddiqui *et al.*, 1989). Conflicting results have been reported regarding the effect of the X gene product on the transcriptional activities of the HBV promoters (Spandau & Lee, 1988; Twu & Robinson, 1989; Siddiqui *et al.*, 1989; Colgrove *et al.*, 1989). Many of the large surface antigen promoter constructs contain the complete X gene and its regulatory sequences (Fig. VII. 6), which gives them the potential to make the X gene product. The possibility that the constructs used in the large surface antigen promoter functional studies were being affected by the presence of the X gene was examined. The full-length PS(1)pLUC construct, which contains the elements necessary for the expression of the X gene, was modified in two ways to interrupt the coding region of the gene (see section VI. A). The plasmid PS(1)pX1LUC was produced by deleting nucleotides 1375 to 1378, which include the initiation codon of the X gene at the *Nco*I site located at nucleotide 1374 in the HBV genome. The plasmid PS(1)pX2LUC contains an 11-nucleotide insertion, creating a frame-shift in the middle of the X open reading frame, produced by inserting a *Sac*I linker into the *Rsr*II site at nucleotide 1577. The absolute and relative activities of the PS(1) promoter, the SV40 early promoter used to direct the expression of the CAT internal control, and the metallothionein promoter used to direct the expression of CAT as an alternative internal control were determined (Fig. VII. 5) in cotransfection experiments performed in Hep3B and HepG2.1 cells. In both cell lines examined, the absolute number of light units produced by the luciferase present in the PS(1)pX1LUC and PS(1)pX2LUC extracts was within 1.5-fold of the number of light units produced by the extract from the unmodified PS(1)pLUC, with the exception of the X1 modification in Hep3B cells, which was three-fold lower than PS(1)pLUC. These uncorrected values indicate that the X gene does not have a large effect, if any, on the large surface antigen promoter in this system. The three-fold effect observed with the X1 construct in the Hep3B cell line was paralleled by a lower CAT activity in that extract, suggesting the possibility that the DNA precipitate may not have transfected as efficiently as the others. The absolute percentage of acetylation by the pMTCAT and pSV2CAT constructs did not vary widely in either cell line. The pMTCAT activities



Hep3B

Transfected DNA	LUC (Light units)	CAT (% acetylation)	LUC/CAT ratio	Relative activity
PS(1)pLUC/pMTCAT	1714	9.7	177	1.00
PS(1)pX1LUC/pMTCAT	1322	11.9	111	0.63
PS(1)pX2LUC/pMTCAT	1787	13.8	129	0.73
PS(1)pLUC/pSV2CAT	3167	12.8	247	1.00
PS(1)pX1LUC/pSV2CAT	880	5.6	157	0.64
PS(1)pX2LUC/pSV2CAT	2123	9.3	228	0.92

HepG2.1

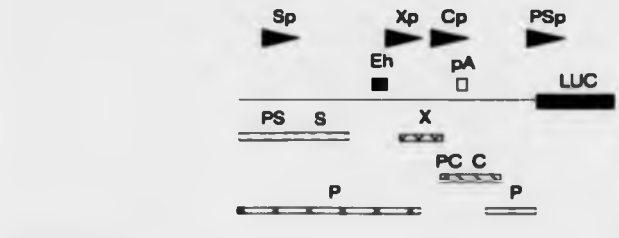
Transfected DNA	LUC (Light units)	CAT (% acetylation)	LUC/CAT ratio	Relative activity
PS(1)pLUC/pMTCAT	335	3.5	95.7	1.00
PS(1)pX1LUC/pMTCAT	386	10.3	37.5	0.39
PS(1)pX2LUC/pMTCAT	236	5.8	40.7	0.43
PS(1)pLUC/pSV2CAT	373	13.7	27.2	1.00
PS(1)pX1LUC/pSV2CAT	427	26.5	16.1	0.59
PS(1)pX2LUC/pSV2CAT	264	34.9	7.6	0.28

Figure VII.5. Influence of the X gene open reading frame on the expression from the large surface antigen, SV40 early and mouse metallothionein promoters in Hep3B and HepG2.1 cells. 15 μg of a luciferase (LUC) construct were cotransfected with 1.5 μg of a chloramphenicol acetyltransferase (CAT) construct into Hep3B and HepG2.1 cells (see Transfected DNA column). Absolute luciferase and CAT values are reported as light units and % acetylation, respectively. The activities are derived from the LUC/CAT ratio and are reported relative to that of the full-length PS(1)pLUC construct with each internal control construct in each cell line. Arrows indicate the position and direction of transcription from the HBV X gene (Xp), core (Cp), PS(1) (PSp) and major surface antigen (Sp) promoters. Boxes indicate the positions of the HBV enhancer I sequence (Eh), HBV polyadenylation recognition sequence (pA), the presurface antigen ORF (PS), surface antigen ORF (S), X gene ORF (X), precore ORF (PC), core ORF (C), polymerase ORF (P), and the luciferase (LUC) ORF. The horizontal lines indicate the HBV sequences present in the various luciferase constructs. The arrowheads designated X1 and X2 indicate the location of the 4-nucleotide deletion and 11-nucleotide insertion in the PS(1)pX1LUC and PS(1)pX2LUC constructs, respectively.

were within two- to three-fold of each other, and the pSV2CAT activities were within approximately two-fold. This result suggests that interpretation of our data using pSV2CAT as an internal control will not be affected by the presence or absence of the coding capacity for the X gene in the experimental constructs. The relative activities calculated from the luciferase values normalized by the CAT values also demonstrated that the X gene does not appear to have a substantial effect on the large surface antigen promoter in this experimental system.

VII. E. Characterization of the Large Surface Antigen Promoter Elements by a 5' Deletion Analysis

A 5' deletion analysis of the HBV genome was performed to determine the regions of the genome important for transcriptional activity from the large surface antigen promoter. A series of 5' deletions of the HBV genome was tested by doing transient transfection experiments using one dedifferentiated hepatoma and three differentiated hepatoma cell lines (Fig. VII. 6). This analysis was performed in several cell lines to examine the possibility that the regulation of this promoter was different in differentiated versus dedifferentiated hepatoma cell lines. All of the constructs tested contained the sequences from -25 to +35 (nucleotides 2784 to 2843) relative to the transcriptional start site at nucleotide 2809. The normalized luciferase activities were reported relative to the activity of the full-length PS(1) promoter construct in each cell line. In the differentiated hepatoma cell lines Hep3B, PLC/PRF/5 (Alexander cells), and HepG2 the effect of the deletions was similar. Essentially full activity of the PS(1) promoter (equal to that produced in the context of the complete HBV genome) was retained when all of the sequences upstream of -90, relative to a transcriptional initiation site at nucleotide 2809 (see construct PS(1)pΔ2840-2718LUC), were deleted (Fig. VII. 6 and 7). Deletion of another 15 nucleotides, through -76 (see construct PS(1)pΔ2840-2733LUC), resulted in the loss of the majority of the PS(1) promoter activity in the differentiated hepatoma cell lines. Deletion of an additional 34



	Relative luciferase activity			
	Hep3B	PLC/PRF/5	HepG2	HepG2.1
PS(1)pLUC	1.00	1.00	1.00	1.00
PS(1)pΔ2840-129LUC	0.84	2.32	1.51	0.58
PS(1)pΔ2840-827LUC	1.61	6.05	2.68	2.05
PS(1)pΔ2840-1070LUC	3.55	8.30	4.60	2.44
PS(1)pΔ2840-1116LUC	1.44	4.18	2.32	0.92
PS(1)pΔ2840-1238LUC	0.28	3.04	3.67	3.93
PS(1)pΔ2840-1374LUC	0.28	4.30	2.36	0.80
PS(1)pΔ2840-1402LUC	0.24	2.81	2.52	0.83
PS(1)pΔ2840-2143LUC	0.14	2.44	0.40	0.21
PS(1)pΔ2840-2425LUC	0.41	0.62	1.69	0.35
PS(1)pΔ2840-2547LUC	0.56	5.31	1.76	0.15
PS(1)pΔ2840-2612LUC	0.42	1.55	2.50	0.64
PS(1)pΔ2840-2707LUC	1.11	3.83	2.64	0.58
PS(1)pΔ2840-2718LUC	1.06	0.70	1.40	0.67
PS(1)pΔ2840-2733LUC	0.05	0.13	0.04	0.84
PS(1)pΔ2840-2734LUC	0.02	0.12	0.14	0.36
PS(1)pΔ2840-2767LUC	0.01	0.04	0.01	0.14
PS(1)pΔ2840-2783LUC	0.00	0.00	0.00	0.01
p19DLUC	0.00	0.00	0.00	0.00

Figure VII.6. Deletion analysis of the HBV large surface antigen promoter in differentiated [Hep3B, PLC/PRF/5 (Alexander), HepG2 and Huh7] and dedifferentiated (HepG2.1) hepatoma cell lines. Arrows indicate the position and direction of transcription from the HBV major surface antigen (Sp), X gene (Xp), core (Cp), and PS(1) (PSp) promoters. Boxes indicate the positions of the HBV enhancer I sequence (Eh), HBV polyadenylation recognition sequence (pA), the presurface antigen ORF (PS), surface antigen ORF (S), X gene ORF (X), precore ORF (PC), core ORF (C) and polymerase ORF (P), and the luciferase (LUC) ORF. The horizontal lines indicate the HBV sequences present in the various PS(1)pΔLUC series plasmids. The plasmid PS(1)pLUC contains the HBV sequences from nucleotides 2840 to 3182/1 to 2843' (nucleotide sequences are designated by using coordinates derived from the Genbank genetic sequence data bank). The designation 2843' has been used to indicate that the nucleotides 2840 to 2843 are present twice in this plasmid and that the nucleotides 2840 to 2843 and 2840' to 2843' are distal and proximal to the LUC ORF, respectively. The HBV sequences deleted from the various plasmids are designated by nucleotide coordinates. The internal control used to correct for transfection efficiencies was pSV2CAT. The activities were calculated as in Fig. VI. 4 and reported relative to the activity of the full-length promoter [PS(1)pLUC] in each cell line.

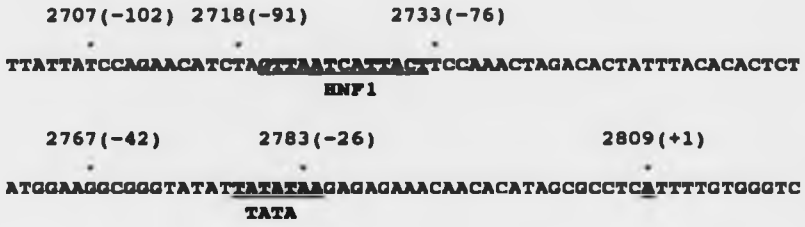


Figure VII.7. Sequence of the HBV large surface antigen promoter region (subtype *ayw*). The numbered nucleotides (•) indicate the breakpoints of the functionally significant deletions. The nucleotide sequences are designated by using coordinates derived from the GenBank genetic sequence data bank. The numbers in parentheses indicate the location of the breakpoints relative to the HBV large surface antigen transcription initiation site at nucleotide 2809 (+1) which is designated by underlining the nucleotide. The underlined sequences represent the hepatocyte nuclear factor 1 (HNF1) binding site and TATA box sequence homologies and the transcription initiation site (+1).

nucleotides, through -42, resulted in almost complete loss of PS(1) promoter activity (see construct PS(1)pΔ2840-2767LUC). These results indicated that a minimal promoter element exists between nucleotides -90 and +35 (coordinates 2719 to 2843), and that an element critical for the activity of the promoter in differentiated hepatoma cell lines lies between nucleotides -90 and -76. The sequences upstream of -90 do not appear to be necessary for maximal activity from the PS(1) promoter, but they may subtly modulate the activity of the promoter. Although individual points in the deletion series differed somewhat, the general level of PS(1) promoter activity remained similar to that of the full-length construct through the deletions to -91 (see construct PS(1)pΔ2840-2718LUC). The variation between individual constructs may represent some level of regulation of the large surface antigen promoter but the effects do not indicate the presence of a critical regulatory element in this region. In Hep3B cells, it appears that the HBV enhancer I (coordinates 1117 to 1238) may have a small effect on the activity of the PS(1) promoter. Deletion of sequences between -1740 and -1693 (see constructs PS(1)pΔ2840-1116LUC and PS(1)pΔ2840-1238LUC) resulted in an approximately four-fold loss of activity. This effect was not observed in any of the other cell lines examined and may reflect a difference between the composition of transcription factors present in the hepatoma cell lines examined.

The full-length PS(1) promoter plasmid [PS(1)pLUC] contains the regulatory elements of all of the HBV promoters. The possibility that the presence of these elements may interfere with transcription from the large surface antigen was examined. Deletion of sequences upstream of -2681 included the major surface antigen promoter and had no effect on the large surface antigen promoter in any of the cell lines examined (see construct PS(1)pΔ2840-129LUC). Similarly, deletion of the X gene promoter region (see constructs PS(1)pΔ2840-1238LUC and PS(1)pΔ2840-1374LUC) and the nucleocapsid promoter region, which includes HBV enhancer II, (see constructs PS(1)pΔ2840-1402LUC and PS(1)pΔ2840-2143LUC) did not have a substantial effect on the activity of the PS(1) promoter. The results of the 5' deletion analysis indicated that transcriptional interference of the large surface antigen promoter from the HBV major

surface antigen, X gene, and nucleocapsid promoters does not occur in this system. Additionally, deletion of the coding region of the X gene did not greatly affect the activity of the PS(1) promoter, supporting the results observed using the X coding region deletion and insertion mutations (Fig. VII. 5).

The region between -90 and -76, which appears critical for PS(1) promoter activity in the differentiated hepatoma cell lines, contains a sequence (GTTAATCATTACT) homologous to the consensus sequence (GTTAATNATTAAC) for the binding site of the liver-enriched transcription factor, hepatocyte nuclear factor 1 (HNF1) (Fig. VII. 7) (Courtois *et al.*, 1988; Lichtsteiner & Schibler, 1989; Frain *et al.*, 1989). HNF1 has been shown to be involved in the regulation of several liver-specific genes (Maire *et al.*, 1989; Hardon *et al.*, 1988; Courtois *et al.*, 1987; Monaci *et al.*, 1988; Sawadaishi *et al.*, 1988), and has been shown to bind this sequence (GTTAATCATTACT) in the HBV large surface antigen promoter (Courtois *et al.*, 1988). Deletion of this region resulted in an 85% to 95% loss in promoter activity, suggesting that HNF1 or a related polypeptide may be responsible for the majority of the large surface antigen promoter activity in differentiated hepatoma cell lines. HNF1 may also be responsible for the preferential activity of the large surface antigen promoter in the differentiated hepatoma cell lines (Table VII. 2). In contrast, the PS(1) promoter activity was not affected by the deletion of the HNF1 recognition sequence in the dedifferentiated hepatoma cell line, HepG2.1. In this cell line, the PS(1) promoter retained full activity when sequences upstream of -75 (see construct PS(1)pΔ2840-2733LUC) were deleted, indicating that a minimal promoter element located between -75 and +35 is important for transcriptional activity in this dedifferentiated hepatoma cell line. Some loss of PS(1) promoter activity in HepG2.1 cells was observed when sequences through -42 (see construct PS(1)pΔ2840-2767LUC) were deleted, but 14% of the activity remained. Essentially complete loss of promoter activity was only observed when sequences including most of the TATA box element at -31 to -25 (see construct PS(1)pΔ2840-2783LUC) were deleted. The lack of dependence on the HNF1 recognition sequence for promoter activity in HepG2.1 cells, and the lower relative activity of the

PS(1) promoter in this cell line suggest that HepG2.1 cells may lack HNF1, or have much lower levels of a functional HNF1 polypeptide than differentiated hepatoma cell lines. These results support the characterization of the HepG2.1 cells as a dedifferentiated line of hepatoma cells.

VII. F. Influence of HNF1 on Transcription from the Four HBV Promoters

The analysis of the transcriptional activity from the large surface antigen promoter in several hepatoma cell lines suggested that the transcription factor HNF1 may be an important regulatory component of the large surface antigen gene. To examine this possibility and the possibility that HNF1 may also regulate transcription from the other HBV promoters, cotransfection experiments were performed using an expression vector encoding the rat HNF1 cDNA (see section VI. A). The expression vector pMTHNF1 or a negative control expression vector pMT was cotransfected into the differentiated hepatoma cell line Huh7 or the dedifferentiated hepatoma cell line HepG2.1 with the full-length HBV promoter-LUC constructs SpLUC, XpLUC, CpLUC, and PS(1)pLUC. The differentiated hepatoma cell line Huh7 was used because it is one of the cell lines in which HBV replication and particle production can occur (Sureau *et al.*, 1986; Tsurimoto *et al.*, 1987; Sells *et al.*, 1987; Chang *et al.*, 1987; Yaginuma *et al.*, 1987) and may represent the tissue culture system closest to the natural environment for the life cycle of HBV, the liver cell. The differentiated hepatoma cell line HepG2 has also been shown to support HBV replication (Sells *et al.*, 1987) and is used in some of the analyses although this cell line exhibits some instability in cell culture. The dedifferentiated hepatoma cell line HepG2.1 was also used because the deletion analysis in this cell line (Fig. VII. 6) suggested that functional HNF1 may be absent from HepG2.1 cells, and these cells would therefore represent a system in which the influence of exogenously expressed HNF1 might be easily measured.

The relative activities of the four HBV promoters were compared in the absence or presence of exogenously expressed HNF1 (Table VII. 3). The activities of the

Table VII. 3. Relative activities[a] of the HBV promoters in HepG2.1 and Huh7 cells in the presence (+HNF1) or absence (-HNF1) of exogenously expressed HNF1 polypeptide[b].

Promoter	Relative Activity					
	HepG2.1			Huh7		
	-HNF1	+HNF1	Fold Induction	-HNF1	+HNF1	Fold Induction
SpLUC	1.00	0.59	0.6	1.00	0.79	0.8
XpLUC	0.83	0.62	0.8	4.49	3.25	0.7
CpLUC	0.18	0.15	0.8	3.91	3.59	0.9
PS(1)pLUC	0.003	0.03	10.0	0.04	0.27	6.8

[a]The activities of the HBV promoters are reported relative to the activity of the major surface antigen promoter (Sp), in the absence of exogenously expressed HNF1, in each cell line. Xp, X gene promoter; Cp, nucleocapsid promoter; PS(1)p, large surface antigen promoter.

[b]The expression vectors used were pMTHNF1 (+HNF1) and pMT (-HNF1).

promoters were reported relative to the activity of the S promoter in the absence of exogenously expressed HNF1 in each cell line. The activity of the PS(1) promoter increased ten-fold in the presence of HNF1 in the dedifferentiated hepatoma cell line, HepG2.1, whereas none of the other HBV promoters was affected by the coexpression of HNF1. In the differentiated Huh7 cells, the PS(1) promoter activity increased seven-fold in the presence of exogenously expressed HNF1, and the activities of the other HBV promoters were not influenced. It appears that HNF1 is not involved in the regulation of surface antigen, X gene, or nucleocapsid promoters of HBV, either directly or indirectly by transcriptional interference, but can influence the activity of the large surface antigen promoter. Increasing the amount of HNF1 expression vector cotransfected in these experiments increased the level of PS(1) promoter activity but did not affect the activity of the other HBV promoters (data not shown). The lack of influence of HNF1 on the S promoter, X promoter, and C promoter may not be surprising as most of the genes regulated by HNF1 contain the recognition sequence within 50 to 200 nucleotides upstream of the transcriptional start site (Mendel & Crabtree, 1991), and in the HBV surface antigen, X gene and core promoter constructs, the HNF1 binding site is located 427, 1760, and 2255 nucleotides upstream of the start of transcription, respectively. In the HBV PS(1) promoter, the HNF1 recognition sequence is 77 nucleotides upstream of the start of transcription.

VII. G. Transactivation of the Large Surface Antigen Promoter is Mediated through the HNF1 Binding Site

Large surface antigen promoter deletion constructs were cotransfected with the pMT or pMTHNF1 expression vector to determine whether the activation of large surface antigen gene transcription by the HNF1 polypeptide was mediated through the binding site located between nucleotides -89 to -77 relative to the start of transcription at nucleotide 2809. The transcriptional activities were reported relative to the activity of the full-length PS(1) promoter in the absence of exogenously expressed HNF1

polypeptide (Fig. VII. 8). Transcription from the constructs PS(1)pLUC and PS(1)pΔ2840-2718LUC (containing sequences -90 to +35) was transactivated in the presence of the pMTHNF1 expression vector by 12- to 15-fold in HepG2.1 cells (Fig. VII. 8A). In contrast, the transcriptional activity from the construct PS(1)pΔ2840-2733LUC, containing sequences -75 to +35, which does not include the HNF1 recognition sequence, was not influenced by the exogenous expression of the HNF1 polypeptide. The results were similar when these experiments were performed in Huh7 cells, although the magnitude of the transactivation was only five- to six-fold (Fig. VII. 8B). In HepG2 cells the transactivation was lower (Fig. VII. 8B), possibly due to the level of endogenous functional HNF1 present in this cell line (Raney *et al.*, 1990). These results strongly suggested that the influence of the HNF1 polypeptide on the large surface antigen promoter is mediated through the HNF1 recognition sequence (GTTAATCATTACT) between nucleotides -90 and -76. These results are also consistent with the observation that HNF1 binds to this sequence in the large surface antigen promoter (Courtois *et al.*, 1988).

VII. H. Gel Mobility Shift Analysis Demonstrates Protein-DNA Complex Using the HBV PS(1) Promoter HNF1 Recognition Sequence

To demonstrate directly that the HNF1 polypeptide produced by expression of the rat cDNA was binding the PS(1) promoter HNF1 recognition sequence, and to determine whether a protein present in the cell lines used in the functional studies could bind the HNF1 recognition sequence, gel mobility shift analysis was performed using nuclear extracts from Huh7, HepG2.1, and HeLa S3 cells and from HepG2.1 cells transfected with the HNF1 expression vector (see section VI. E for preparation of extracts and VI. F for gel mobility shift protocol). A double-stranded oligonucleotide containing the HBV PS(1) promoter HNF1 binding site was radiolabelled and incubated with the extracts before electrophoresis on a 4% polyacrylamide gel. A gel retardation product was observed when 2 μg of Huh7 nuclear extract was incubated with the

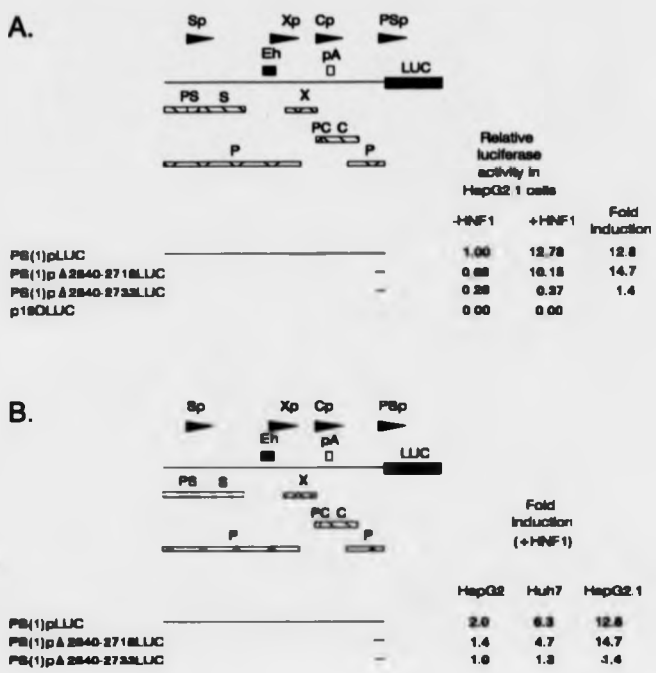


Figure VII.8. **(A) Influence of the HNF1 polypeptide exogenously expressed in HepG2.1 cells on the transcription from various large surface antigen promoter constructs, which included [PS(1)pLUC and PS(1)pΔ2840-2718LUC] or did not include the HNF1 binding site [PS(1)pΔ2840-2733LUC]. The construct p19DLUC was transfected as a negative control. The relative activities of the large surface antigen promoter constructs were determined in the presence (+HNF1) and the absence (-HNF1) of exogenously expressed HNF1 polypeptide using the expression vectors pMTHNF1 and pMT, respectively. The activities were calculated as in Fig. VI. 4 and reported relative to the activity of the full-length promoter [PS(1)pLUC] in the absence of exogenously expressed HNF1 polypeptide. Arrows indicate the positions and directions of transcription from the HBV surface (Sp), X gene (Xp), core (Cp), and PS(1) (PSp) promoters, respectively. Boxes indicate the positions of the HBV enhancer I sequence (Eh), HBV polyadenylation sequence (pA), presurface antigen ORF (PS), surface antigen ORF (S), X gene ORF (X), precore ORF (PC), core ORF (C), polymerase ORF (P), and luciferase ORF (LUC). The horizontal lines indicate the HBV sequences present in the various PS(1) promoter plasmids. The HBV sequences deleted from the various plasmids are designated by nucleotide coordinates derived from the GenBank genetic sequence data bank. The internal control used to correct for transfection efficiencies was pSV2CAT. (B) Influence of the expression of the HNF1 polypeptide on the transcriptional activity from the large surface antigen promoter in HepG2, Huh7, and HepG2.1 cells.**

radiolabelled HNF1 double-stranded oligonucleotide (Fig. VII. 9), suggesting that a polypeptide which recognized the HNF1 binding site was present in the Huh7 nuclear extracts. This protein-DNA complex was competed when excess (20-fold, 200-fold, and 2000-fold molar excess) unlabelled HNF1 double-stranded oligonucleotide was added to the reaction, indicating that the protein in the complex was binding specifically to the HNF1 oligonucleotide. A similar complex with a slightly greater mobility was observed when 12 μg of HepG2.1 or 12 μg of HeLa S3 nuclear extract were incubated with the radiolabelled HNF1 oligonucleotide. The complex formed with the HepG2.1 extract was efficiently competed by the same amounts of unlabelled HNF1 oligonucleotide that were used to compete the Huh7 complex, suggesting that the protein in this complex was also specifically binding to the HNF1 recognition sequence. The unlabelled HNF1 oligonucleotide did not compete as efficiently for the HeLa S3 complex as for the Huh7 and HepG2.1 complexes, but complete competition was achieved at a 2000-fold molar excess of oligonucleotide. The HeLa S3 cells may contain a polypeptide with similar migration properties to the polypeptides that bind to the HNF1 recognition sequence in Huh7 and HepG2.1 cell extracts, but the specificity of binding appears lower than that observed with the hepatoma cell extracts. Gel mobility shift analysis also demonstrated that nuclear extracts made from Hep3B cells contained a polypeptide, that bound the radiolabelled HNF1 oligonucleotide, which appeared to be similar in size and concentration to the polypeptide in Huh7 nuclear extracts that bound the HNF1 recognition sequence (data not shown). Similar amounts of HepG2.1 and HeLa S3 nuclear extract displayed binding to the HNF1 oligonucleotide at much lower levels than in the Huh7 and Hep3B extracts (data not shown). The relative intensities of the bands in lanes 2, 6 and 10, which contain 2 μg of Huh7, 12 μg of HepG2.1 and 12 μg of HeLa nuclear extract, respectively, suggested that the concentration of a polypeptide binding the HNF1 recognition sequence was approximately six-fold greater in Huh7 cells than in HepG2.1 or HeLa cells. The complex formed with the Huh7 extract appeared to comprise a closely migrating doublet, and the complex formed with the HepG2.1 and HeLa S3 extracts migrated at the position of the lower band in the Huh7 and Hep3B

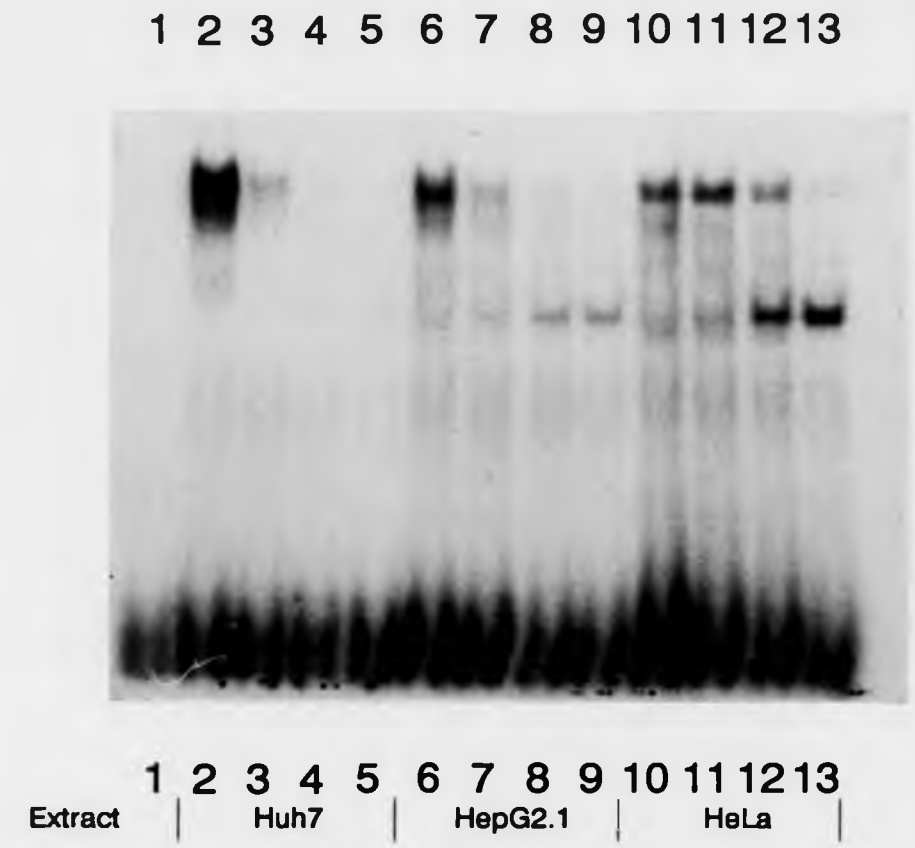


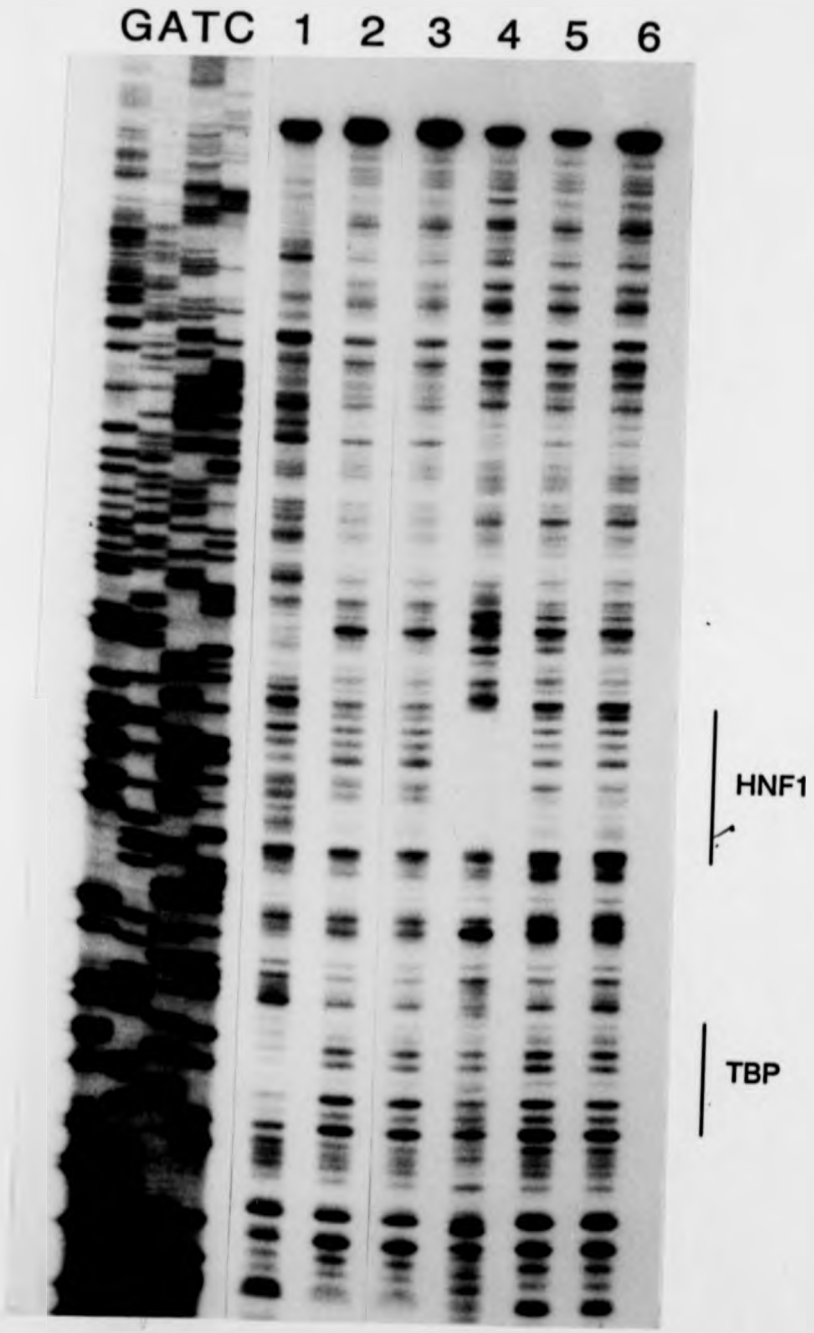
Figure VII.9. Gel retardation analysis of the HBV large surface antigen promoter HNF1 binding region. The ^{32}P-labelled double-stranded oligonucleotide probe containing the HNF1 recognition sequence located in the PS(1) promoter was analyzed for its ability to form complexes with DNA-binding proteins present in Huh7, HepG2.1 and HeLa cell nuclear extracts. Unlabelled double-stranded HNF1 oligonucleotide was used as competitor DNA to demonstrate the specificity of the observed complexes. Two μg Huh7 nuclear extract (lanes 2-5), 12 μg HepG2.1 nuclear extract (lanes 6-9) and 12 μg HeLa nuclear extract (lanes 10-13) were incubated with the radiolabelled double-stranded HNF1 oligonucleotide probe. Unlabelled double-stranded HNF1 oligonucleotide was also added to the reactions shown in lanes 3, 7, 11 (20-fold molar excess), 4, 8, 12 (200-fold molar excess) and 5, 9, 13 (2000-fold molar excess). Lane 1 is a control in which the HNF1 oligonucleotide was incubated with buffer instead of nuclear extract.

doublet. One possible explanation for this observation is that the differentiated hepatoma cell lines may contain the HNF1 polypeptide as well as a related polypeptide, vHNF1 or HNF1β, which is a slightly smaller polypeptide and also binds the HNF1 recognition sequence (Baumhueter *et al.*, 1988; Cereghini *et al.*, 1988), and that the dedifferentiated hepatoma cell line HepG2.1 and the non-hepatoma cell line HeLa S3 express only the variant of HNF1, vHNF1, or an unrelated polypeptide, which forms a complex migrating in the position of the lower band of the doublet in the Huh7 extract. The complex formed using the Huh7 nuclear extracts migrated in the same position as the gel retardation product using HepG2.1 cells expressing the HNF1 cDNA (Fig. VII. 16). The untransfected HepG2.1 cells expressed a polypeptide which formed a complex of a slightly greater mobility than that of the HNF1 cDNA transfected cells. These observations are consistent with the suggestion that Huh7 cells express a functional HNF1 polypeptide and HepG2.1 cells express a related polypeptide at a level too low to activate the PS(1) promoter through the HNF1 binding site.

VII. I. DNase I Footprinting Analysis of the HBV PS(1) Promoter using Nuclear Extracts and Purified TATA Binding Protein

DNase I footprint analysis demonstrated that purified TATA binding protein (TBP) can bind the PS(1) promoter TATA box element located from -31 to -25 relative to the start of transcription (see section VI. G for DNase I footprinting protocol). The nucleotides on a 423-base-pair radiolabelled PS(1) promoter fragment that were protected from DNase I digestion by purified TBP spanned -25 to -40 on the long (minus) strand, encompassing the sequence TTATATAATATACCCG (Fig. VII. 10). The DNase I footprinting analysis also identified a binding site for a protein which was present in Huh7 nuclear extracts. Nuclear extracts prepared from HepG2.1 and HeLa S3 cells did not protect the same region (Fig. VII. 10). The region of the PS(1) promoter fragment protected from DNase I digestion by Huh7 nuclear extract spanned nucleotides -69 to -99 on the minus strand, encompassing the sequence

Figure VII.10. DNase I footprinting analysis of the long (minus) strand of the HBV large surface antigen promoter. The 423-nucleotide HBV DNA fragment from -379 to +44 was 5'-end labelled at +44 and incubated in the presence (lane 1) or absence (lane 2) of 4 units of purified TBP (TFIID, Promega), or in the presence (lanes 4, 5 and 6, respectively) or absence (lane 3) of 50 μg Huh7, HepG2.1 or HeLa nuclear extract. An unrelated DNA sequence (GATC) is adjacent to lane 1 and was used as a size standard to localize the TBP and HNF1 binding sites.


GATC 1 2 3 4 5 6
HNF1
TBP

AGTTTGGAAGTAATGATTAACTAGATGTTCT. This protected region included the 13-nucleotide sequence containing homology to the consensus recognition sequence for the transcription factor HNF1 (AGTAATGATTAAC). Consistent with the liver-enriched distribution of HNF1 (Courtois *et al.*, 1988; Courtois *et al.*, 1987; Rey-Campos *et al.*, 1991; Cereghini *et al.*, 1990; Baumhueter *et al.*, 1988) and the functional data indicating the importance of the HNF1 binding site for PS(1) promoter activity in differentiated hepatoma cell lines (Fig. VII. 6 and 8), it seems likely that the protein binding the HNF1 recognition site in the differentiated hepatoma cell line Huh7 is HNF1 or a related polypeptide. The absence of a detectable footprint over the HNF1 recognition sequence using the HepG2.1 and HeLa S3 nuclear extracts is also consistent with the suggestion that the dedifferentiated HepG2.1 cell line and the non-hepatoma HeLa S3 cell line lack a sufficient level of functional HNF1 polypeptide to transactivate the large surface antigen promoter.

VII. J. Influence of Orientation of the HNF1 Binding Site on Transcriptional Transactivation by the HNF1 Polypeptide

The HNF1 recognition sequence, GTTAATNATTAAC, is an inverted palindrome to which the HNF1 polypeptide has been shown to bind as a dimer (Frain *et al.*, 1989). In light of this information, it was of interest to determine whether the orientation of the HNF1 binding site affected the level of transcriptional transactivation contributed by the HNF1 polypeptide. A double-stranded oligonucleotide containing the large surface antigen HNF1 binding site was inserted upstream of the HBV sequences in the minimal promoter construct PS(1)pΔ2840-2767LUC which contained promoter sequences -41 to +35 (see section VI. A and Fig. VII. 7). The plasmid PS(1)pΔ2840-2767HNF1(+)LUC contained the HNF1 binding site in the same orientation as it was found in the HBV genome, and the recognition sequence was located 22 nucleotides upstream of the PS(1) promoter TATA box element (located between nucleotides -31 and -25; Fig. VII. 7) The plasmid PS(1)pΔ2840-2767HNF1(-)LUC contained the HNF1 binding site in the opposite

orientation to that found in the HBV genome, and the recognition sequence was 21 nucleotides upstream of the PS(1) promoter TATA box element. In the plasmids PS(1)pLUC and PS(1)pΔ2840-2718LUC, the HNF1 binding site and TATA box element were separated by 45 nucleotides, as in the HBV genome. The constructs containing the synthetic HNF1 binding site oligonucleotides were transfected into HepG2.1 cells in the absence (-HNF1) or presence (+HNF1) of the expression vector pMT or pMTHNF1, respectively. The activities were reported relative to the activity of the full-length PS(1) promoter in the absence of the HNF1 expression vector (Fig. VII. 11). The relative activity of the PS(1) minimal promoter construct lacking the HNF1 site (plasmid PS(1)pΔ2840-2767LUC) was 12-fold lower than the full-length plasmid, PS(1)pLUC. When the synthetic HNF1 oligonucleotide was inserted into the minimal promoter construct (constructs PS(1)pΔ2840-2767HNF1(+)LUC and PS(1)pΔ2840-2767HNF1(-)LUC), the relative activity was nine- to ten-fold higher than the activity of the positive control, PS(1)pLUC. The presence of the HNF1 binding site, in either orientation, appeared to be responsible for a substantial increase in transcriptional activity from the large surface antigen minimal promoter construct, PS(1)pΔ2840-2767LUC. One possible explanation for the ten-fold higher relative activity of the constructs containing the synthetic HNF1 binding site compared with the full-length PS(1) promoter may be that the closer positioning of the HNF1 recognition sequence to the TATA box results in a more efficient utilization of an endogenous polypeptide present in HepG2.1 cells which binds the HNF1 site. Experiments examining the effects of spacing between the HNF1 binding site and TATA box have been performed and are described in section VII. M. From the results described above, it appeared that the sequences deleted from these minimal promoter constructs (from -76 to -42) were not necessary for maximal promoter activity when the HNF1 binding site was present. In fact, the deletion of these sequences resulted in higher relative activity from the synthetic promoter (constructs PS(1)pΔ2840-2767HNF1(+)LUC or PS(1)pΔ2840-2767HNF1(-)LUC) than from the minimal promoter construct containing these sequences and the HNF1 binding site (construct PS(1)pΔ2840-2718LUC). In addition to the

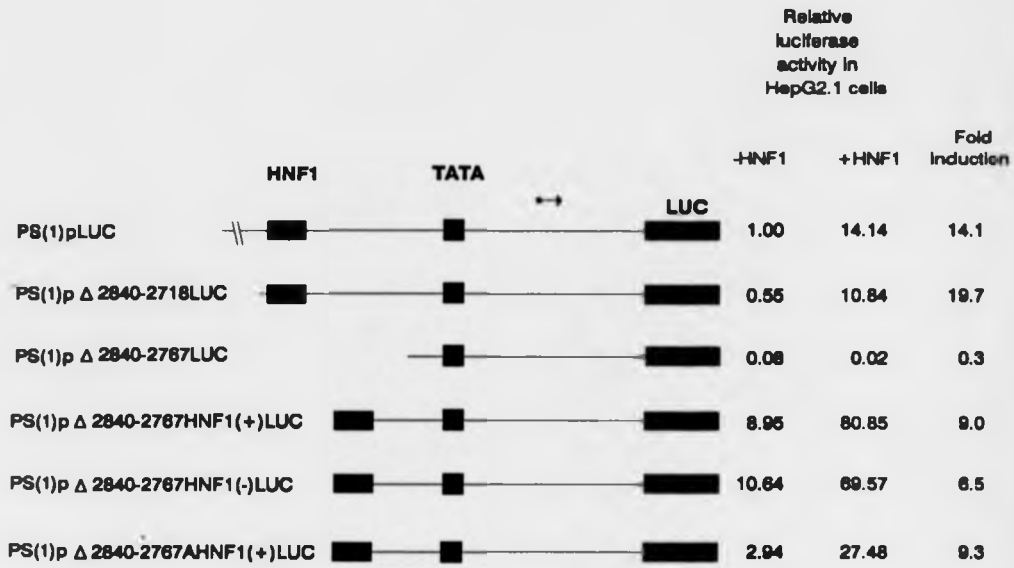


Figure VII.11. Influence of the orientation and sequence of the HNF1 binding site on the transcriptional activity from large surface antigen minimal promoter constructs containing single synthetic HNF1 binding sites in HepG2.1 cells. The arrow indicates the location of the large surface antigen transcription initiation site. The boxes indicate the positions of the HNF1 binding site, the HBV TATA box element, and the luciferase (LUC) ORF. The HBV sequences deleted from the various plasmids are designated by nucleotide coordinates derived from the GenBank genetic sequence data bank. The plasmid PS(1)pΔ2840-2767HNF1(+)LUC contains a synthetic large surface antigen HNF1 binding site in the same orientation as it is found in the HBV genome. The plasmid PS(1)pΔ2840-2767HNF1(-)LUC contains a synthetic large surface antigen HNF1 binding site in the opposite orientation to that found in the HBV genome. The plasmid PS(1)pΔ2840-2767AHNF1(+)LUC contains a synthetic HNF1 binding site from the human albumin promoter in the same orientation relative to the direction of transcription as is found in the albumin gene. The relative activities of the minimal promoter constructs were determined in the presence (+HNF1) and the absence (-HNF1) of exogenously expressed HNF1 polypeptide using the expression vectors pMTHNF1 and pMT, respectively. The activities were calculated as in Fig. VI. 4 and reported relative to the activity of the full-length promoter [PS(1)pLUC] in the absence of exogenously expressed HNF1. The internal control used to correct for transfection efficiencies was pSV2CAT.

possibility that spacing between the promoter elements may affect the relative activity from the promoter, there exists the possibility that the sequences between -76 and -42 may comprise a negative element which, when deleted, results in a relative increase in promoter activity. It is also possible that the construction of the synthetic promoter elements produced an artificial positive element at the junction of the oligonucleotide sequence and the HBV minimal promoter sequence. It appears that the orientation of the HNF1 binding site does not affect the level of transcriptional activity from these constructs as the relative activity from the (+) and (-) orientation constructions was equivalent. When the HNF1 expression vector was included in the transfection (+HNF1), the constructs containing the synthetic HNF1 binding site were inducible to a similar degree as the full-length promoter construct. These results indicate that the large surface antigen HNF1 binding site inserted into the PS(1) promoter minimal construct was able to mediate transcriptional transactivation by the HNF1 polypeptide regardless of its orientation, as may have been expected, as HNF1 is known to bind its recognition sequence as a dimer (Frain *et al.*, 1989; Nicosia *et al.*, 1990; Chouard *et al.*, 1990).

VII. K. Influence of Sequence of the HNF1 Binding Site on Transcriptional Transactivation by the HNF1 Polypeptide

HNF1 has been reported to interact with a number of liver-specific genes (Cereghini *et al.*, 1988; Cereghini *et al.*, 1990; Courtois *et al.*, 1988; Courtois *et al.*, 1987; Feuerman *et al.*, 1989; Hardon *et al.*, 1988; Lichtsteiner & Schibler, 1989; Maire *et al.*, 1989; Ryffel *et al.*, 1989; Tsutsumi *et al.*, 1989; Vaulont *et al.*, 1989), and it has been shown to bind its recognition sequence both as a homodimer and as a heterodimer with the related polypeptide, vHNF1 (Chouard *et al.*, 1990; Nicosia *et al.*, 1990; De Simone *et al.*, 1991; Rey-Campos *et al.*, 1991). An analysis using the human albumin promoter HNF1 binding site was reported (Nicosia *et al.*, 1990) which mapped the transcriptional activation domains of HNF1 to a different region of the polypeptide from this analysis using the HBV PS(1) promoter HNF1 binding site (see section VII. O; Fig. VII. 15). To

determine whether this observed difference were due to the sequences of the HNF1 binding sites, a double-stranded oligonucleotide containing the human albumin promoter HNF1 binding site was cloned into the HBV large surface antigen minimal promoter construct PS(1)pΔ2840-2767LUC to create PS(1)pΔ2840-2767AHNF1(+)LUC. It was necessary to determine whether the full-length HNF1 polypeptide could activate transcription from the promoter construct containing the albumin HNF1 recognition sequence. The sequence of the human albumin promoter HNF1 binding site (GTTAATAATCTAC) differs from the HBV PS(1) promoter HNF1 binding site (GTTAATCATTACT) at four of the 12 conserved positions in the consensus HNF1 binding site (GTTAATNATTAAC) (Courtois *et al.*, 1988). Each of these HNF1 binding sites differs from the consensus recognition sequence by only two nucleotides in the 3' half of the sequence. The relative transcriptional activity from the plasmid containing the synthetic albumin HNF1 binding site, PS(1)pΔ2840-2767AHNF1(+)LUC, was most similar to that of the PS(1) promoter, both in the absence (-HNF1) and presence (+HNF1) of exogenously expressed HNF1 polypeptide (Fig. VII. 11). This result demonstrated that the human albumin HNF1 binding site mediated transcriptional activation by HNF1 polypeptide exogenously expressed in HepG2.1 cells to a similar extent as the HBV PS(1) promoter HNF1 binding site contained in a minimal promoter construct. The sequence differences between the albumin and HBV HNF1 binding sites did not appear to influence the ability of the recognition sequence to mediate the induction of transcriptional activation by the HNF1 polypeptide in this system.

VII. L. Transcriptional Activity from the PS(1) Promoter HNF1 Binding Site and TATA Box Element

It appeared that the HNF1 binding site was critical for the majority of the large surface antigen promoter activity in differentiated hepatoma cell lines. In order to determine whether this promoter element together with the TATA box element could support transcriptional activity, synthetic promoter constructs containing the PS(1)

promoter TATA box element, or the PS(1) promoter HNF1 binding site, or both (see section VI. A) upstream of the luciferase ORF were tested for their transcriptional activities in HepG2.1 cells in the absence (-HNF1) or presence (+HNF1) of the HNF1 expression vector (Fig. VII. 12). The activities were reported relative to the activity of the full-length PS(1) promoter construct in the absence of exogenously expressed HNF1 (-HNF1). The construct containing only the TATA box element upstream of the luciferase ORF (construct pHBVTATALUC) exhibited transcriptional activity of approximately one-fifth the level of the PS(1) promoter and it was not inducible by expression of HNF1. The construct containing only the PS(1) promoter HNF1 binding site upstream of the luciferase ORF (construct pHNF1LUC) produced transcriptional activity approximately two-fold lower than the full-length PS(1) promoter and this activity was inducible to a similar extent as the PS(1) promoter. The construct containing both the HNF1 binding site and the TATA box element (construct pHNF1TATALUC) displayed a level of transcriptional activity equivalent to the full-length PS(1) promoter, suggesting that these two elements may be responsible for the majority of the activity from the large surface antigen promoter, and that the other HBV sequences are dispensable for the activity observed in this assay system. The transcriptional activity of this construct was also increased ten-fold by the expression of the HNF1 polypeptide, similar to the induction of the full-length promoter. These results suggested that regulated transcriptional activity similar to that observed for the complete HBV PS(1) promoter can be achieved using a synthetic promoter comprising two promoter elements, the HNF1 recognition sequence and the TATA box element. It is possible that subtle regulation by additional transcription factors may influence the level of activity from the large surface antigen promoter to a degree that is not detectable in these experiments. It appears, however, that the interaction of HNF1 with the PS(1) promoter is a very important facet of the regulation of the large surface antigen gene.

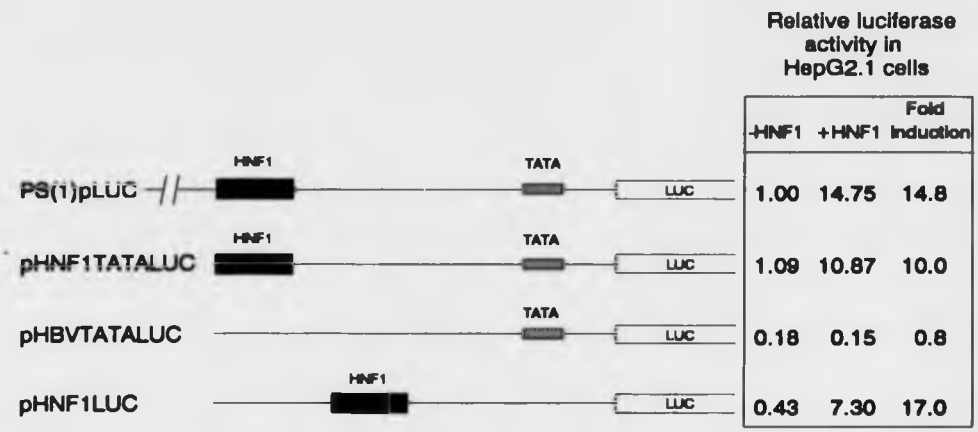


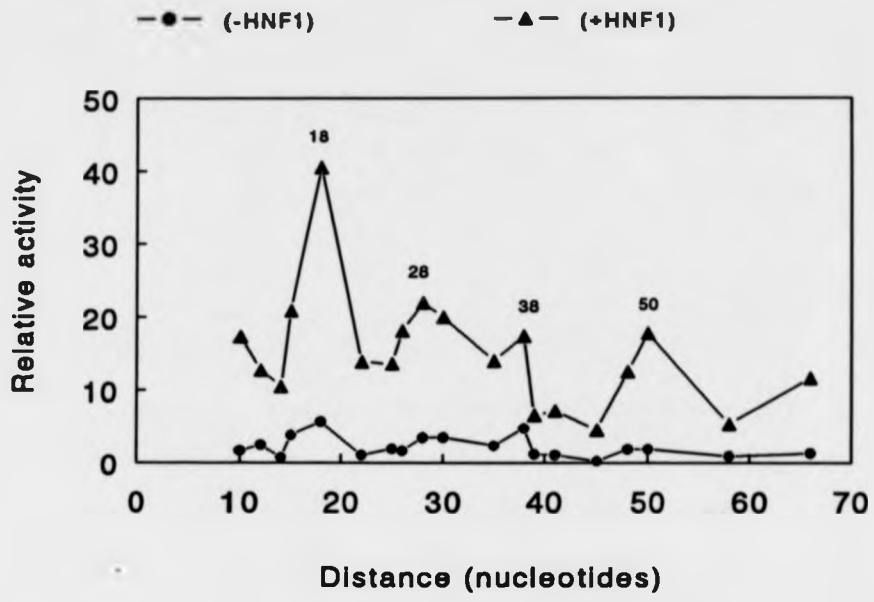
Figure VII.12 Transcriptional activities from constructs containing the PS(1) promoter HNF1 binding site and TATA box element. The relative activities of the constructs were determined in HepG2.1 cells in the presence (+HNF1) or absence (-HNF1) of exogenously expressed HNF1 polypeptide using the expression vectors pMTHNF1 and pMT, respectively. The activities were calculated as in Fig. VI. 4 and reported relative to the activity of the full-length promoter [PS(1)pLUC] in the absence of exogenously expressed HNF1 polypeptide. The internal control used to correct for transfection efficiencies was pSV2CAT. The HBV sequences of plasmid PS(1)pLUC are represented by a thin line. The locations of the HNF1 and TATA box sequence homologies in all of the constructs are indicated by boxes labelled HNF1 and TATA, respectively. The luciferase gene sequences are indicated by the box labelled LUC. In the plasmids pHNF1TATALUC, pHBVTATALUC and pHNF1LUC, linker sequences are represented by thin lines.

VII. M. Influence of Spacing between the HNF1 Binding Site and the TATA Box Element of the HBV PS(1) Promoter

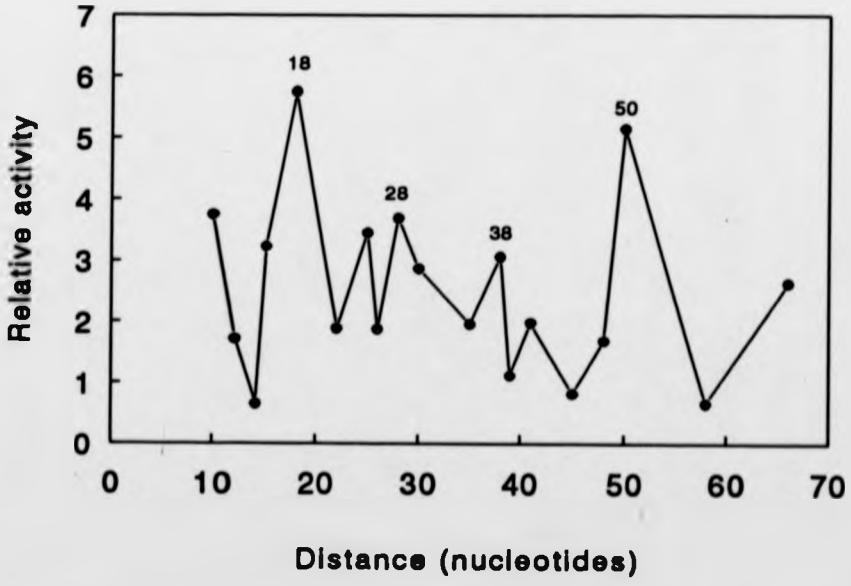
In the synthetic construct pHNF1TATALUC, the number of nucleotides between the two promoter elements is 39 nucleotides, whereas in the HBV PS(1) promoter, there are 45 nucleotides between these two elements. The results shown in Fig. VII. 12 suggested that a small change in the spacing between the HNF1 binding site and the TATA box element does not affect the activity or the inducibility of the construct. However, the sequences between the two elements differ in the two constructs being compared [PS(1)pLUC and pHNF1TATALUC] and these could affect the level of transcription from the plasmid. In addition, in a previous experiment in which the HNF1 binding site was inserted upstream of a minimal promoter construct (see section VII. J, construct PS(1)pΔ2840-2767HNF1(+)LUC), creating a construct in which the spacing between the HNF1 binding site and the TATA box element was 22 nucleotides, the basal level of transcription was approximately ten-fold higher than that of the full-length promoter (Fig. VII. 11). The experiments in Fig. VII. 12 showed that transcriptional activity and HNF1 responsiveness could be achieved using a synthetic promoter containing only the HNF1 binding site and the TATA box element of the PS(1) promoter. To examine the effect of changing the spacing between these two elements, a series of plasmids was constructed from pHNF1TATALUC in which the spacing between the HNF1 binding site and the TATA box element was changed by the addition or deletion of nucleotides (see section VI. A). The spacing varied from 10 to 66 nucleotides between the two binding sites. These constructs were transfected into HepG2.1 cells and their transcriptional activities in the absence (-HNF1) and presence (+HNF1) of exogenously expressed HNF1 were compared with the activity from the full-length PS(1) promoter in the absence (-HNF1) of exogenously expressed HNF1 (Fig. VII. 13A). The results of this transfection series indicated that the spacing between the HNF1 recognition sequence and the TATA box element does not have a large effect on the activity from these synthetic promoters when transfected in HepG2.1 cells in the

Figure VII.13 The influence of spacing between the HNF1 binding site and the TATA box of the PS(1) promoter in (A) HepG2.1 cells and (B) Huh7 cells. Synthetic promoter constructs containing the PS(1) promoter HNF1 recognition sequence and the TATA box element separated by a range of nucleotide distances were transfected into HepG2.1 cells in the absence or presence of exogenously expressed HNF1 polypeptide (A), or into Huh7 cells in the absence of exogenously expressed HNF1 polypeptide (B). The relative activities were calculated as in Fig. VI. 4 and reported relative to the activity of the full-length PS(1)pLUC construct in the absence of exogenously expressed HNF1 polypeptide. The internal control used to correct for transfection efficiencies was pSV2CAT. The distance (nucleotides) indicates the number of nucleotides between the 3' and 5' boundaries of the HNF1 and TATA binding sequences, respectively. The constructs which generated peaks of relative activity are indicated by the distance between the binding sites (18, 28, 38, 50).

A.


(-HNF1)
(+HNF1)
Relative activity
Distance (nucleotides)
18
28
38
50

B.


Relative activity
Distance (nucleotides)
18
28
38
50

absence of HNF1. Although some fluctuation of activity between individual points seemed to occur, and two high values occurred at a distance of 18 and 38 nucleotides, the level of activity did not change enough to indicate that a particular distance between the two binding sites was highly beneficial or detrimental for the transcriptional activity in HepG2.1 cells. In the presence of HNF1 polypeptide, however, relatively high levels of activity occurred at distances of 18, 28, 38 and 50 nucleotides between the HNF1 and TATA binding sites. The apparent 10-nucleotide periodicity of higher relative transcriptional activities suggested that the binding of HNF1 and TATA binding protein (TBP) on the same face of the DNA increased the efficiency of transcriptional transactivation of these plasmids. The activity of each of the constructs was induced five- to 15-fold by HNF1. Similar results were obtained when the experiment was performed in Huh7 cells in the presence of the endogenous HNF1 polypeptide (Fig. VII. 13B). Relatively higher levels of activity were observed when the distance between the HNF1 and TATA box recognition sequences were 18, 28, 38 and 50 nucleotides. The relative activities from the synthetic constructs in which the distances between the HNF1 and TATA elements were 22 and 45 nucleotides were approximately the same as those from the corresponding constructs containing HBV sequences, PS(1)pΔ2840-2767HNF1(+)LUC and PS(1)pLUC, respectively (data not shown). The data also seemed to suggest a slight reduction in transcriptional activity as the distance between the binding sites increased, although additional constructs with greater distances between these sites would need to be tested to verify this trend.

VII. N. Clustered Point Mutational Analysis of the Minimal PS(1) Promoter

From the data described above, it appeared that the spacing between the HNF1 recognition sequence and the TATA box element could be changed without loss of transcriptional activity or inducibility by the HNF1 polypeptide (see section VII. M and Fig. VII. 13). It also appeared that the sequences between the HNF1 binding site and TATA box element were not important for these activities, as the synthetic spacing

constructs contained linker sequences, rather than the endogenous HBV sequences, between the two elements. In order to determine whether the HBV sequences did play some role in the regulation of transcription from a minimal promoter without altering the spacing between the HNF1 binding site and TATA box element, clustered point mutations were made in the region of the promoter from the HNF1 binding site to the 5' end of the TATA box element (see section VI. A). Each clustered mutation changed approximately ten nucleotides (Fig. VII. 14A). These constructs were cotransfected into HepG2.1 cells in the presence of the negative control vector pMT (-HNF1) or the HNF1 expression vector pMTHNF1 (+HNF1), and their relative activities were compared with that of the full-length PS(1) promoter construct and the HBV constructs from which the mutant constructs were made (see section VI. A). None of the four clustered point mutations M1 to M4 appeared to have a large influence on the activity from the promoter as the relative activities (-HNF1) varied only from 1.1- to 2.6-fold over the full-length promoter (Fig. VII. 14B). The mutations M1 to M4 (see constructs PS(1)pM1LUC to PS(1)pM4LUC) were constructed from the plasmid PS(1)pΔ2840-2707LUC and contained mutations of the sequences between the HNF1 binding site and the TATA box element. The transcriptional activities of the mutant promoters relative to the activity observed with the PS(1)pΔ2840-2707LUC clone were 0.4 to 0.9 (-HNF1), indicating that the mutated sequences were not critical for promoter activity. In the presence of exogenously expressed HNF1 (+HNF1), the M1 to M4 mutant constructs exhibited activities similar to those of the full-length promoter and their parental construct, PS(1)pΔ2840-2707LUC, and their levels of induction were also similar. The mutant constructs were induced to levels between the levels of the PS(1)pΔ2840-2707LUC cnstruct and the full-length clone in these experiments, indicating that the M1 to M4 mutations do not appear to affect the activity of the promoter, in the absence or presence of exogenously expressed HNF1 in HepG2.1 cells. These constructs were also transfected into Huh7 cells in the absence of the HNF1 expression vector to determine the effects of the mutations in the presence of endogenous HNF1. The results of those transfection experiments (Fig. VII. 14B) demonstrated that the M3 and M4

Figure VII.14 Clustered point mutational analysis of the minimal PS(1) promoter. (A) The sequence (subtype *ayw*) of the HBV PS(1) promoter region from -89 to +1 is shown. Coordinates of the large surface antigen promoter region are derived from the Gen Bank genetic sequence data bank and their position relative to the transcription initiation site (+1) is shown in parentheses. The regions of homology to the HNF1 recognition sequence and the TATA box element, and the start of transcription are underlined. The sequence of the clustered point mutation is shown for each mutant construct. The nucleotides which differ from the HBV sequence are indicated in lowercase; unchanged nucleotides are indicated in uppercase. The M1, M2, M3, M4 mutations were introduced into the PS(1)pΔ2840-2707LUC construct and the M5 mutation was introduced into the PS(1)pΔ2840-2425LUC construct.
(B) Relative activities of the clustered point mutation constructs in HepG2.1 cells in the presence (+HNF1) or absence (-HNF1) of exogenously expressed HNF1 polypeptide, and in Huh7 cells in the absence of exogenously expressed HNF1 polypeptide. The activities were calculated as in Fig. VI. 4 and reported relative to the activity of the full-length promoter [PS(1)pLUC] in the absence of exogenously expressed HNF1 polypeptide. The internal control used to correct for transfection efficiencies was pSV2CAT. The diagrammatic representation of the construct PS(1)pLUC is shown at the top. The HBV sequences of the plasmids are represented by a thin line. The locations of the HNF1 and TATA box sequence homologies are indicated by boxes labelled HNF1 and TATA, respectively. The luciferase gene sequences are indicated by the box labelled LUC. The locations of the M1, M2, M3, M4 and M5 mutations are indicated by shaded boxes.

A.

B.

	Relative Luciferase Activity			
	HepG2.1			Huh7
	-HNF1	+HNF1	Fold Induction	
PS(1)pLUC	1.00	8.43	8.4	1.00
PS(1)pΔ2840-2425LUC	0.44	3.59	8.2	0.16
PS(1)pΔ2840-2707LUC	2.75	9.40	3.4	1.38
PS(1)pΔ2840-2733LUC	0.13	0.16	1.2	0.00
PS(1)pM1LUC	2.56	7.65	3.0	0.26
PS(1)pM2LUC	1.12	6.32	5.6	0.25
PS(1)pM3LUC	2.20	10.08	4.6	0.98
PS(1)pM4LUC	1.62	7.30	4.5	0.81
PS(1)pΔ2840-2425M5LUC	0.36	0.28	0.8	0.00

mutations of the sequences between the HNF1 and TATA box binding sites did not affect the activity from the full-length promoter or the minimal promoter construct PS(1)pΔ2840-2707LUC in the presence of the endogenous HNF1 in Huh7 cells. The M1 and M2 mutations, which lie 3 to 12 nucleotides and 13 to 22 nucleotides, respectively, downstream of the HNF1 binding site appeared to have a small effect, four- to five-fold, on the activity from the promoter constructs PS(1)pLUC and PS(1)pΔ2840-2707LUC. It is possible that the proximity to the HNF1 site of these mutations could affect the interaction between HNF1 and its recognition sequence. However, these mutated constructs do retain considerable transcriptional activity, unlike the constructs that do not contain the HNF1 binding site (see construct PS(1)pΔ2840-2733LUC and PS(1)pΔ2840-2425M5LUC).

The 5' deletion analysis of the large surface antigen promoter provided evidence that the HNF1 recognition sequence was important for the majority of the transcriptional activity from this promoter in the presence of endogenous (in differentiated hepatoma cell lines) or exogenously expressed (cotransfected into dedifferentiated HepG2.1 cells) HNF1 polypeptide (see sections VII. E and G; Fig. VII. 6 and 8). To examine the possibility that other HBV sequences could compensate for the HNF1 recognition sequence in its absence, the mutant construct PS(1)pΔ2840-2425M5LUC was made. In this plasmid the sequence of the HNF1 binding site was mutated in the presence of the promoter sequences from -383 to +35 [plasmid PS(1)pΔ2840-2425LUC]. It appeared (Fig. VII. 14) that the HNF1 recognition sequence did not contribute to the activity from the PS(1) promoter in the dedifferentiated hepatoma cell line HepG2.1 in the absence of exogenously produced HNF1 polypeptide [see column (-HNF1)] because the HNF1 binding site mutant construct retained essentially full activity relative to its parental construct (compare construct PS(1)pΔ2840-2425M5LUC with PS(1)pΔ2840-2425LUC). These results are consistent with the deletion analysis (Fig. VII. 6) and with the suggestion that the HepG2.1 cell line lacks a sufficient level of functional HNF1 polypeptide (or related polypeptide) to transactivate the HBV large surface antigen promoter. The construct containing the mutated HNF1 binding site

was not transactivatable in HepG2.1 cells in the presence of exogenously expressed HNF1 polypeptide [see column (+HNF1)], suggesting that other HBV sequences were unable to compensate for the HNF1 recognition sequence by mediating transactivation of the promoter by the HNF1 polypeptide. These results are consistent with the suggestion that the HNF1 polypeptide binds specifically to its recognition sequence in the HBV large surface antigen promoter to mediate its effect. In Huh7 cells, the relative activity of the HNF1 mutant construct PS(1)pΔ2840-2425M5LUC was undetectable in these experiments. This observation is consistent with the data from the deletion analysis (Fig. VII. 6 and 8) demonstrating the importance of the HNF1 binding site for transcriptional activity from the PS(1) promoter in differentiated hepatoma cell lines. These data suggest that the HNF1 recognition sequence is necessary to mediate transcriptional activation of the large surface antigen promoter by the HNF1 polypeptide.

VII. O. Characterization of the Transcriptional Activation Domain of HNF1

In order to determine the transcriptional activation domain of the HNF1 polypeptide, a series of deletion mutants of the cDNA in the HNF1 expression vector was constructed (see section VI. A). A functional analysis of these deletion plasmids was performed by cotransfecting the deletion series into HepG2.1 cells with plasmids containing the HNF1 binding site to measure their ability to transactivate transcription from the promoter constructs containing the HNF1 recognition sequence. A 3' and internal deletion analysis of the cDNA was performed because the dimerization and DNA binding domains of the polypeptide had previously been identified in the amino-terminal 281 amino acids of the polypeptide (Chouard *et al.*, 1990; Frain *et al.*, 1989; Nicosia *et al.*, 1990). The fold-induction of the promoter constructs by the truncated HNF1 polypeptides were reported relative to the activity of the promoter construct in the absence of the HNF1 expression vector (in the presence of the negative control expression vector pMT) (Fig. VII. 15). The HNF1 polypeptide designations indicate the amino acids which are not present in the product of the cDNA deletion expression

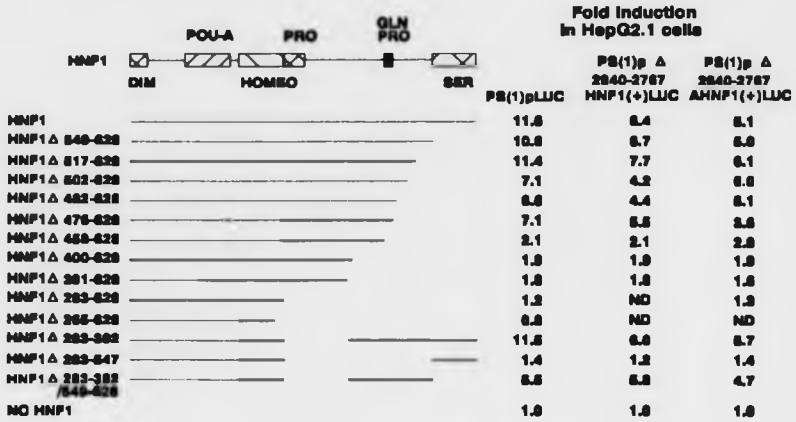


Figure VII.15. Deletion analysis of the HNF1 polypeptide and identification of the domain involved in the transcriptional activation of the large surface antigen promoter [PS(1)pLUC] and minimal promoter constructs containing synthetic HNF1 binding sites [PS(1)pΔ2840-2767HNF1(+)LUC and PS(1)pΔ2840-2767AHNF1(+)LUC] in HepG2.1 cells. The plasmid PS(1)pΔ2840-2767HNF1(+)LUC contains a synthetic large surface antigen HNF1 binding site in the same orientation as it is found in the HBV genome. The plasmid PS(1)pΔ2840-2767AHNF1(+)LUC contains a synthetic HNF1 binding site from the human albumin promoter in the same orientation relative to the direction of transcription as is found in the albumin gene. DIM, dimerization domain; POU-A, region of POU subdomain A homology; HOMEO, homeobox-like domain; PRO, proline-rich transcriptional activation domain; GLU/PRO, glutamine- and proline-rich transcriptional activation domain; SER, serine-rich transcriptional activation domain. The horizontal lines indicate the HNF1 polypeptide amino acid residues present in the various truncated HNF1 polypeptides expressed in HepG2.1 cells. The HNF1 polypeptide designations indicate the amino acids which are not present in the product of the cDNA deletion expression vectors. The HNF1 designation represents the full-length HNF1 polypeptide. The activities were calculated as in Fig. VI. 4 and reported relative to the activity of each reporter construct in the absence of exogenously expressed HNF1 (NO HNF1). ND; not done.

vector. The deletion series was tested on three promoter constructs: the full-length PS(1) promoter [PS(1)pLUC]; the PS(1) minimal promoter construct containing the HNF1 binding site [PS(1)pΔ2840-2767HNF1(+)LUC]; and the PS(1) minimal promoter construct containing the human albumin HNF1 binding site [PS(1)pΔ2840-2767AHNF1(+)LUC]. The induction of the PS(1) promoter by the full-length HNF1 expressed in HepG2.1 cells was approximately 11-fold. The results indicated that the carboxy-terminal 112 amino acids were not necessary for maximal levels of induction of the large surface antigen promoter (see HNF1Δ517-628 and column PS(1)pLUC). The region between amino acids 502 and 516 (see HNF1Δ502-628 and HNF1Δ517-628) appeared to contribute to the maximal level of transactivation by the HNF1 polypeptide as deletion of the nucleotides encoding these residues resulted in a loss of approximately 35% of the transactivation of the PS(1) promoter. No further reduction in transactivation was observed as a result of deletion of an additional 26 amino acids (see HNF1Δ482-628 and HNF1Δ476-628). The region between amino acids 458 and 475, however, appeared essential for the majority of the transactivation of the PS(1) promoter, as deletion of the cDNA through the region encoding these residues resulted in almost complete loss of induction by the truncated HNF1 polypeptide (see HNF1Δ458-628 and HNF1Δ476-628). Although the absolute level of induction of the construct PS(1)pΔ2840-2767HNF1(+)LUC by the full-length HNF1 polypeptide was a little lower than that of the PS(1)pLUC construct, the region of the polypeptide necessary to transactivate the minimal promoter construct containing the HBV HNF1 binding site was clearly identified by the deletion analysis as the same as the transactivation domain identified using the PS(1) promoter construct [see column PS(1)pΔ2840-2767HNF1(+)LUC]. This 18-amino-acid region contains five glutamine and four proline residues. Other transcription factor activation domains also have a high concentration of particular amino acids. For example, the transactivation domain of Sp1 is rich in glutamine residues (Courey & Tjian, 1988) and the transactivation domain of CTF is rich in proline residues (Mermod *et al.*, 1989).

The transactivation domain identified by this analysis differs from the previously identified transactivation domains of HNF1. An *in vitro* transcription analysis performed

with a promoter containing the human albumin HNF1 binding site identified two regions of the HNF1 polypeptide necessary for transcriptional activation (Nicosia *et al.*, 1990). These were the regions from amino acids 281 to 318 and from 547 to 628. To determine whether this observed difference in the mapping of the transactivation domains were due to the different HNF1 recognition sequences, the deletion analysis was performed on the minimal promoter construct containing the human albumin HNF1 binding site [Fig. VII. 15, column PS(1)pΔ2840-2767AHNF1(+)LUC]. The level of induction of this promoter construct by the HNF1 polypeptide was equivalent to the level of induction of the PS(1) minimal promoter construct containing the HBV HNF1 binding site, consistent with previous results showing that the sequence of the HNF1 recognition site did not influence the effect of transactivation in the context of the PS(1) minimal promoter construct (see section VII. K. and Fig. VII. 11). Deletion of the region encoding the activation domain previously identified by *in vitro* transcription analysis in the carboxyl terminus of the polypeptide (amino acid residues 547 to 628) (Nicosia *et al.*, 1990) did not influence the induction of the albumin HNF1 binding site-containing promoter in this system (see construct HNF1Δ549-628 and column PS(1)pΔ2840-2767AHNF1(+)LUC). Deletion of sequences encoding the other activation domain identified by *in vitro* transcription analysis (amino acid residues 281 to 318) also had no effect on the activation of the promoter construct containing the albumin HNF1 binding site in this analysis (see constructs HNF1Δ391-628 and HNF1Δ283-628). The region which appeared to mediate the HNF1 transactivation was contained between amino acids 400 (HNF1Δ400-628) and 481 (HNF1Δ482-628), the region that contains the transactivation domain identified using the PS(1) promoter in transfection assays (amino acids 458 to 475). This is distinct from the transactivation domains previously identified using the albumin promoter HNF1 binding site in an *in vitro* transcription analysis (Nicosia *et al.*, 1990). The difference in the mapping of the transcriptional activation domains may be due to the differences between an *in vitro* transcription analysis and a transient transfection analysis. These results indicate that the same transactivation domain of the HNF1 polypeptide can mediate the effect of HNF1 through two different HNF1 binding

sites in the context of the minimal promoter construct used in this system.

Three internal deletion constructs were tested to confirm the results of the carboxy-terminal deletions in the mapping of the transactivation domains of the HNF1 polypeptide. Deletion of the region encompassing the transactivation domain identified by the carboxy-terminal deletions (HNF1Δ283-547) resulted in essentially complete loss of transactivation of all of the promoters tested. In contrast, the HNF1 truncated polypeptides in which the regions encompassing the previously identified transactivation domains of HNF1 (HNF1Δ283-392 and HNF1Δ283-392/549-628) were deleted were able to transactivate all of the promoters examined. The results of the 3' and internal deletion analyses are consistent with the identification of a transactivation domain of the HNF1 polypeptide located in a region between amino acids 393 to 516, distinct from the activation domains identified in the *in vitro* transcription analysis using the human albumin HNF1 binding sites (Nicosia *et al.*, 1990). These results indicated that the previously identified activation domains are not necessary for induction of the promoters containing the large surface antigen and human albumin HNF1 binding sites used in these experiments.

VII. P. Gel Mobility Shift Analysis of Truncated HNF1 Polypeptides

To demonstrate that the exogenously expressed truncated HNF1 polypeptides were able to bind the HNF1 recognition sequence, gel mobility shift analysis was performed using nuclear extracts prepared from HepG2.1 cells transfected with the truncated HNF1 expression vectors (Fig. VII. 16). Nuclear extract prepared from untransfected Huh7 cells was also used (lane 2) and shown to produce a gel retardation complex which migrated in the same position as the product of the HepG2.1 cells expressing the full-length HNF1 cDNA (lane 4). This result was consistent with the evidence from the transient transfection analysis that the Huh7 cell line expresses a functional HNF1 polypeptide. The gel retardation complex formed with the untransfected HepG2.1 nuclear extract (lane 3) appeared to be less abundant than the

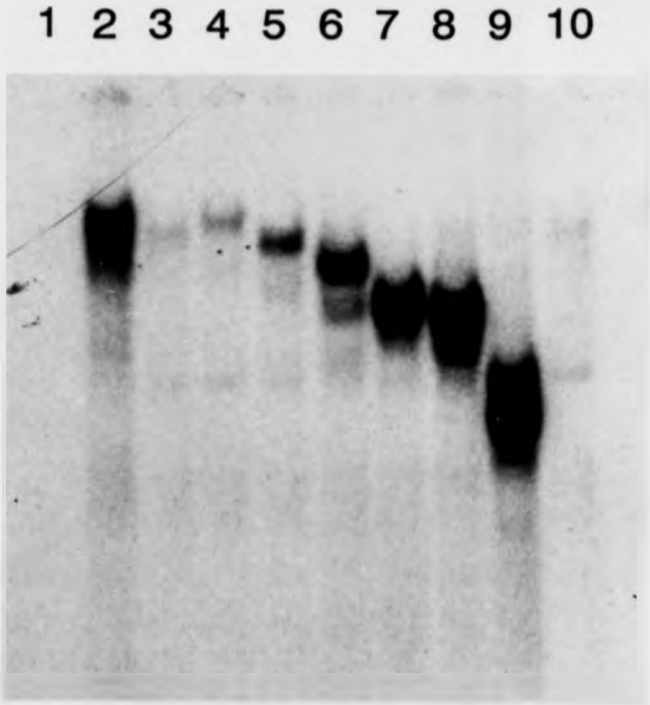


Figure VII.16. Gel retardation analysis of truncated HNF1 polypeptides. Nuclear extracts were prepared from HepG2.1 cells transfected with 30 μg of the expression vector (lane 4) pMTHNF1, (lane 5) pMTHNF1Δ502-628, (lane 6) pMTHNF1Δ458-628, (lane 7) pMTHNF1Δ400-628, (lane 8) pMTHNF1Δ391-628, (lane 9) pMTHNF1Δ283-547, (lane 10) pMT. Nuclear extracts were also prepared from Huh7 cells (lane 2) and HepG2.1 cells (lanes 3-10) as described in section VI.E. Six μg nuclear extract were incubated with the radiolabelled double-stranded HNF1 oligonucleotide as described in section VI. F. Lane 1 is a control in which the HNF1 oligonucleotide was incubated with buffer instead of nuclear extract.

Huh7 complex and migrated slightly faster, consistent with the suggestion that HepG2.1 cells express a low level of an HNF1-related polypeptide, possibly vHNF1, which binds to the PS(1) promoter HNF1 recognition sequence. The truncated HNF1 polypeptides produced from the expression vectors pMTHNF1Δ502-628 (lane 5), pMTHNF1Δ458-628 (lane 6), pMTHNF1Δ400-628 (lane 7), pMTHNF1Δ391-628 (lane 8), and pMTHNF1Δ283-547 (lane 9) produced gel mobility shift complexes with the HNF1 recognition sequence that decreased in size corresponding to the extent of the cDNA deletion. The polypeptide HNF1Δ283-547 (lane 9) generated several gel mobility shift products, suggesting that the polypeptide may have been proteolytically cleaved or that it formed several conformations due to the extent of the deletion. These expression vectors were chosen for this experiment as representatives of expression vectors whose products did transactivate (pMTHNF1, pMTHNF1Δ502-628), or did not transactivate (pMTHNF1Δ458-628, pMTHNF1Δ400-628, pMTHNF1Δ391-628, pMTHNF1Δ283-547) the large surface antigen promoter in HepG2.1 cell transfections. In addition, the product of the expression vector pMTHNF1Δ502-628 exhibited a small decrease in the amount of transactivation it mediated on all of the promoter constructs tested, the product of the vector pMTHNF1Δ458-628 defined the breakpoint for the region critical for transactivation of the large surface antigen promoter HNF1 binding site constructs, and the product of the vector pMTHNF1Δ283-547 confirmed by internal deletion the region important for transactivation of all of the promoter constructs. It was important to determine whether the expected polypeptides were expressed and could bind to the recognition sequence, especially for those that did not mediate transactivation. The sample containing buffer only (lane 1) did not form a complex with the HNF1 binding site. The sample containing extract made from HepG2.1 cells transfected with the negative control expression vector (lane 10) formed the same complex as the untransfected HepG2.1 extract sample. This analysis demonstrated that truncated HNF1 polypeptides were expressed and able to bind the HNF1 recognition sequence.

VII. Q. Effect of Expression Vector Levels on Transcriptional Activation of the PS(1) Promoter in HepG2.1 Cells

The gel mobility shift analysis (Fig. VII. 16) results suggested that transfection of the HNF1 expression vectors may result in the synthesis of different amounts of the truncated HNF1 polypeptides. To determine whether the amount of expression vector transfected affected the transcriptional activation of the large surface antigen promoter, various amounts of HNF1 expression vector plasmids were transfected into HepG2.1 cells and the induction of the PS(1) promoter was examined (Fig. VII. 17). The results demonstrated that the levels of induction of the PS(1) promoter increased as the amounts of the expression vector plasmids pMTHNF1, pMTHNF1Δ549-628, and pMTHNF1Δ283-392/549-628 transfected were increased. Expression vectors which produced truncated HNF1 polypeptides which had little effect on the activity of the PS(1) promoter in the analysis of the functional activation domain of HNF1 (pMTHNF1Δ400-628, pMTHNF1Δ283-547, and pMTHNF1Δ283-628) did not transactivate the large surface antigen promoter when increased amounts (up to 30 μg) of the vector were transfected. This analysis demonstrated that the mapping of the transactivation domains of the HNF1 polypeptide was not affected by the amount of expression vector transfected. The levels of transactivation in this analysis were lower than previously observed when 1.5 μg of expression vector were transfected and may be due to the increased amount of DNA transfected into the cells (46.5 μg total) compared with the standard assay (18 μg total).

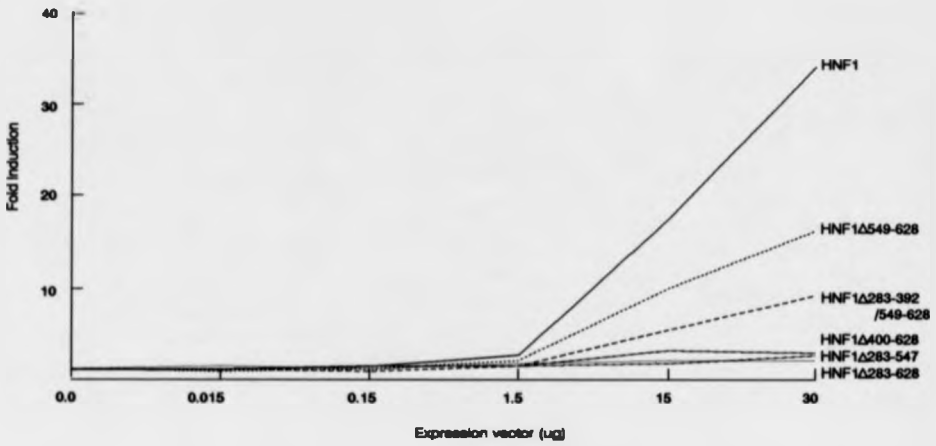


Figure VII.17. Effect of expression vector levels on the transcriptional activation of the large surface antigen promoter in HepG2.1 cells. Cells were transfected with 15 μg of PS(1)pLUC, 1.5 μg of pSV2CAT, and 30 μg of pMT plus expression vector DNA. The indicated amount of expression vector was combined with pMT DNA such that the total amount of DNA was 30 μg. Each transfection mixture therefore contained a total of 46.5 μg DNA. Fold induction indicates the amount of transcriptional activity relative to the activity from PS(1)pLUC in the absence of exogenously expressed HNF1 polypeptide. The HNF1 polypeptide designations indicate the amino acids which are not present in the product of the cDNA deletion expression vectors. The HNF1 designation represents the full-length polypeptide.

VIII. DISCUSSION

Hepatitis B virus is a hepatotropic virus of highly restricted host range, infecting only human beings and chimpanzees. It is known that the major target of HBV infection is the hepatocyte, but the mechanism(s) of this tropism is not understood. Possible restrictions at many steps in the viral life cycle, such as the need for a hepatocyte-specific viral receptor, or for tissue specific transcription factors, may limit infectivity predominantly to hepatocytes. HBV infection of primary hepatocyte cultures has been achieved (Rijntjes *et al.*, 1988; Gripon *et al.*, 1988), but tissue culture cell lines have not been successfully infected. However, transfection of viral DNA into differentiated hepatoma tissue culture cell lines has resulted in the production of replication intermediates and viral particles (Sureau *et al.*, 1986; Tsurimoto *et al.*, 1987; Sells *et al.*, 1987; Chang *et al.*, 1987; Yaginuma *et al.*, 1987). This suggests that at least one factor in the hepatotropism of the virus is related to entry of the virus into the cell. However, it has also been shown that transgenic mice carrying an integrated HBV genome express the 2.1 kb surface antigen and 3.5 kb pregenomic RNAs primarily in the liver and kidney (Araki *et al.*, 1989; Farza *et al.*, 1988), and the enhancers I and II and core promoter regulatory regions appear to display liver cell-type specificity in transfection experiments (Jameel & Siddiqui, 1986; Karpen *et al.*, 1988; Antonucci & Rutter, 1989; Honigwachs *et al.*, 1989; Yee, 1989; Chang & Ting, 1989; Yuh & Ting, 1993), suggesting that tissue-specific transcriptional regulation may also play a role in the tissue specificity of the virus. To investigate the transcriptional regulation of HBV, and of the large surface antigen gene in particular, several tissue culture cell lines were used (Fig. VII. 3). In previously described tissue culture transfection experiments (De-Medina *et al.*, 1988; Pourcel *et al.*, 1982; Siddiqui *et al.*, 1986), the expression of the major surface antigen gene appeared to be less specific than in transgenic mice, which indicates that tissue culture cell lines may be different with respect to the composition of transcription factors from primary cultures or tissues. Huh7 and HepG2 cells are differentiated human hepatoblastoma cell lines which have been shown to produce viral

particles after transfection of HBV DNA (Sureau *et al.*, 1986; Tsurimoto *et al.*, 1987; Sells *et al.*, 1987; Chang *et al.*, 1987; Yaginuma *et al.*, 1987). This indicates that these cell lines have all of the components necessary for the life cycle of the virus and may represent the closest tissue culture system to the *in vivo* environment. The human hepatocellular carcinoma cell lines PLC/PRF/5 (Alexander) and Hep3B contain integrated HBV sequences and have been shown to express major surface antigen (Alexander *et al.*, 1976; Aden *et al.*, 1979; Macnab *et al.*, 1976; Knowles *et al.*, 1980). The PLC/PRF/5 cell line has also been shown to complement surface antigen-minus mutant HBV genomes to produce virus (Okamoto *et al.*, 1993). The dedifferentiated HepG2.1 cell line is a derivative of the HepG2 cell line which arose during continuous culturing of the HepG2 cells (Raney *et al.*, 1990). The morphological and biochemical phenotype of this cell line is distinct from the parental HepG2 line. It does not have the characteristic hepatocyte-like epithelial morphology, and its γ-glutamyltransferase activity is much lower than that of all of the other differentiated hepatoma cell lines. The HeLa S3 line is a human cervical carcinoma cell line, and the NIH 3T3 line is a mouse fibroblast cell line. The dedifferentiated hepatoma and non-hepatoma cell lines have been used in the analysis of transcriptional regulation of the large surface antigen promoter to try to distinguish regulatory components of the virus which may be specific for differentiated hepatoma cell lines capable of supporting viral replication.

The relative strengths of the four hepatitis B virus promoters were examined in several cell lines in the context of the complete viral genome located upstream of a luciferase (LUC) reporter gene (Table VII. 2). As the major surface antigen promoter displayed relatively high transcriptional activity in all cell lines examined (Siddiqui *et al.*, 1986; Pourcel *et al.*, 1982; De-Medina *et al.*, 1988), the activities of the other three HBV promoters were reported relative to the activity of the surface antigen promoter in each cell line to compare the cell type specificity of the promoters. The full length genome was located upstream of the LUC reporter gene such that expression of luciferase was directed by the major surface antigen, X gene, nucleocapsid, or large surface antigen promoter in the constructs SpLUC, XpLUC, CpLUC, and PS(1)pLUC, respectively (Fig.

VII. 4). The HBV DNA sequences were derived from the clone pCP10 (Dubois *et al.*, 1980) of the HBV subtype *ayw*. A clone of this subtype was used because the *ayw* cloned DNA has been shown to produce viral particles which are infectious in chimpanzees (Will *et al.*, 1982; Will *et al.*, 1985), indicating that all of the sequences necessary for the life cycle of the virus have been maintained. The constructs were made with the complete genome to include all of the *cis*-acting regulatory sequences which may influence the activity of the promoters. The analysis of the activities of the four HBV promoters in several cell lines demonstrated that the major surface antigen promoter and the X gene promoter displayed the greatest amount of transcriptional activity in most of the cell lines examined. The nucleocapsid, or core, promoter exhibited similar levels of activity to the surface antigen promoter in the differentiated hepatoma cell lines, but relatively less activity in the dedifferentiated hepatoma and nonhepatoma cell lines. The large surface antigen promoter was the weakest of the four promoters in every cell line tested, but like the core promoter, demonstrated higher relative activity compared with the surface antigen promoter in the differentiated hepatoma cell lines. These results suggested that the nucleocapsid and large surface antigen promoters may be subject to a degree of cell-type specific regulation and therefore may contribute to the hepatotropism of the virus. The differentiated hepatoma cell lines may contain factors which interact specifically with the regulatory regions of the nucleocapsid and large surface antigen genes to increase transcription from these promoters. The relative levels of activity from the major surface antigen, nucleocapsid and large surface antigen promoters were similar to the 2.1 kb, 3.5 kb and 2.4 kb RNA levels observed during HBV infection and in differentiated hepatoma cell lines producing viral particles (Cattaneo *et al.*, 1984; Chang *et al.*, 1987; Su *et al.*, 1989b; Sureau *et al.*, 1986; Tsurimoto *et al.*, 1987; Yaginuma *et al.*, 1987; Yokosuka *et al.*, 1986). These results indicated that the relative activities of the viral promoters appear to be important for virus production, suggesting that coordinate regulation of transcription may be employed by the virus to generate the levels of mRNAs necessary for production of the appropriate levels of viral polypeptides required for assembly of the virus. Large

surface antigen is an essential component of the virion, or Dane particle (Ueda *et al.*, 1991). The relative weakness of the PS(1) promoter suggested that regulation of the level of the 2.4 kb RNA, from which the large surface antigen is translated, may represent an important factor in the viral life cycle. The high relative level of activity from the XpLUC construct in these experiments was not consistent with the low or undetectable level of 0.7 kb RNA observed in infected tissue or cell lines producing viral particles (Cattaneo *et al.*, 1984; Chang *et al.*, 1987; Su *et al.*, 1989b; Sureau *et al.*, 1986; Tsurimoto *et al.*, 1987; Yaginuma *et al.*, 1987; Yokosuka *et al.*, 1986). This observation could be attributed to a difference in the stability of the luciferase mRNA relative to the stability of the 0.7 kb mRNA, or to the fortuitous generation of a positive element in the X promoter as a result of cloning the HBV sequences into the luciferase vector, or to a loss of regulation of the X gene in the transient transfection system.

Many studies have indicated that the product of the X gene can transactivate several viral and cellular promoters, including HBV promoters and the SV40 enhancer and early promoter (Spandau & Lee, 1988; Twu & Schloemer, 1987; Zahm *et al.*, 1988; Seto *et al.*, 1988; Twu & Robinson, 1989; Siddiqui *et al.*, 1989; Colgrove *et al.*, 1989). As the full-length large surface antigen promoter construct PS(1)pLUC contains the complete X gene and the sequences necessary for its expression, the possibility that the X gene product was influencing the activities of the SV40 early promoter, the mouse metallothionein promoter, or the large surface antigen promoter in these transient transfection analyses was examined (section VII. D and Fig. VII. 5). The activities of the full-length PS(1)pLUC construct and two additional constructs derived from PS(1)pLUC and designed to interrupt the coding capacity of the X gene, PS(1)pX1LUC and PS(1)pX2LUC, were analyzed in HepG2.1 and Hep3B cells in the presence of one of two internal control plasmids, pSV2CAT or pMTCAT. The absolute numbers of light units and the corrected relative activities of the PS(1)pLUC construct and the X mutation PS(1)pLUC constructs did not vary more than two- to three-fold in either of the cell lines examined. The CAT activities from the cotransfected internal control pMTCAT did not vary more than two- to three-fold, and the activities from pSV2CAT

were within approximately two-fold. These variations were no greater than those expected as normal variation of these assays. These results indicated that in this system, if the X gene product were being made, it did not influence the activities of the large surface antigen, SV40 early, or mouse metallothionein promoters enough to alter the interpretation of the data. These data were consistent with the results of the 5' deletion analysis (see section VII. E, Fig. VII. 6) which demonstrated that deletion of the X ORF did not substantially affect the activity of the large surface antigen promoter. These results are inconsistent with reports of X transactivation of the SV40 enhancer and early promoter (Spandau & Lee, 1988; Zahm *et al.*, 1988; Twu & Robinson, 1989; Siddiqui *et al.*, 1989). It is possible that the X gene product was not made in this system, and this could account for the lack of effect of the X gene on these promoters observed in this analysis. It is also possible that the amount of X gene product expressed in this system was inappropriate for transactivation of these promoters. It has been shown that the effect of the X gene product on the HBV enhancer is dose dependent (Faktor & Shaul, 1990). An alternative explanation for the lack of effect of the X gene product on these promoters is that the cell lines used for these analyses may lack a factor necessary to mediate the effect of the X polypeptide. The X gene product has not been shown to bind DNA and is thought to act through cellular factors such as CREB, ATF-2, AP-1, AP-2 and NF-kβ which bind DNA and may mediate its effects (Maguire *et al.*, 1991; Faktor & Shaul, 1990; Twu *et al.*, 1989a; Twu *et al.*, 1989b; Siddiqui *et al.*, 1989; Siddiqui *et al.*, 1987; Lucito & Schneider, 1992; Seto *et al.*, 1990). The precise role of the X polypeptide in HBV infection has not been adequately defined and, in fact, the X gene product is not needed for HBV replication in cell culture (Blum *et al.*, 1992) but does appear necessary for replication of the WHV in woodchucks (Chen *et al.*, 1993).

The higher relative activity of the large surface antigen promoter in differentiated hepatoma cell lines suggested that liver cell type specific factors may be interacting with regulatory regions of the large surface antigen gene to mediate the increased activity in these hepatoma cell lines. When this analysis was initiated, the regulation of the large surface antigen promoter had not been well characterized. The

initiation site of the 2.4 kb RNA had been mapped by S1 nuclease and primer extension analyses in cell culture systems to nucleotide 2809 (Sells *et al.*, 1988; Yaginuma *et al.*, 1987), which is 25 to 32 nucleotides downstream of a TATA box element characteristic of eukaryotic promoters. One analysis of transcriptional activity indicated that an HBV fragment containing 400 nucleotides upstream of the 2.4 kb transcription initiation site contained promoter activity (Siddiqui *et al.*, 1986). In an attempt to characterize the *cis*-acting regulatory elements in the HBV genome which mediated the apparent preferential liver cell type activity of the large surface antigen promoter, transient transfection analyses were performed. A 5' deletion analysis of the HBV DNA sequences contained in the large surface antigen promoter-LUC construct PS(1)pLUC was performed in dedifferentiated and differentiated hepatoma cell lines (Fig. VII. 6). The PS(1)pLUC construct contained the complete HBV genome located upstream of the luciferase ORF such that expression of luciferase was directed by the large surface antigen promoter. In the differentiated hepatoma cell line Hep3B, PS(1) promoter activity decreased approximately four-fold when the sequences including the HBV enhancer I were deleted [see construct PS(1)pΔ2840-1238LUC]. This decrease was not observed in any other cell line, suggesting that the Hep3B cells may contain a positive transcription factor binding to the enhancer I region which is not present in the other cell lines examined. These results are not entirely consistent with those of an analysis of the large surface antigen promoter using CAT as a reporter molecule in transfections of human hepatoma cell lines (Chang & Ting, 1989). In that study, a 400 base pair enhancer fragment inserted downstream of the CAT gene, which was under the control of a large surface antigen promoter fragment (-380 to +19), increased the activity from the PS(1) promoter by 30- to 40-fold in Huh7 and HepG2 cells, but did not increase the PS(1) promoter activity in the poorly differentiated hepatoma line HA22T/VGH or in HeLa cells. The effects of the enhancer on the activity of the large surface antigen promoter in that study were much greater than those observed in this analysis and the effects were seen in HepG2 cells and Huh7 cells, in which no influence by the enhancer was detected in this analysis. It is possible that differences in subtype (*adw* versus *ayw*) or differences in the

plasmid constructions used could account for these differences. The plasmid constructs used in this analysis contain the complete HBV genome interrupted only by the insertion of the luciferase vector sequences downstream of the large surface antigen transcription initiation site and the effect of the enhancer was measured by deletion of the enhancer sequences, whereas Chang and Ting (Chang & Ting, 1989) inserted a 400 base pair PS(1) promoter fragment upstream of the CAT gene and measured promoter activity in the presence or absence of a 400 base pair enhancer fragment inserted downstream of the CAT gene. The full-length constructs used in this analysis were designed to maintain all of the sequences which might affect transcription, which could then be systematically deleted or mutated to observe changes in transcriptional activities. This approach was taken to preserve the context of the *cis*-acting elements as closely as possible to the native genome, rather than inserting selected fragments out of context around a reporter molecule. Other analyses have also shown some effect of the enhancer I on PS(1) promoter activity, but these analyses also differ in their constructions and subtypes of HBV sequences used (Shaul *et al.*, 1985; Patel *et al.*, 1989).

All of the cell lines retained transcriptional activity when the sequences upstream of -90 [see construct PS(1)pΔ2840-2718LUC] relative to the transcription initiation site at HBV nucleotide 2809 were deleted. These results indicated that, in this system, the other HBV promoters and enhancer II (located in the nucleocapsid promoter region) elements did not have a measurable influence on the activity of the large surface antigen promoter. These results are consistent with an analysis in which the enhancer II region did not affect the activity of the large surface antigen promoter in Huh7 cell transient transfection experiments (Zhou & Yen, 1990). The minor variation observed in the activities of the deletion constructs upstream of -90 may be the effect of interactions of transcription factors with the HBV genome that subtly modulate the activity from the large surface antigen promoter. If that is so, the influences are small and do not affect the majority of the transcriptional activity. In the differentiated hepatoma cell lines, the transcriptional activity decreased to levels of 4% to 13% of the activity of the full-length promoter (compare constructs PS(1)pLUC with PS(1)pΔ2840-2733LUC) when the

sequences upstream of -75 were deleted. This deletion defined a 15-nucleotide region which appeared to be critical for the activity of the large surface antigen promoter in differentiated hepatoma cell lines. This region contains a sequence (GTTAATCATTACT) homologous to the consensus recognition sequence (GTTAATNATTAAC) for the transcription factor hepatocyte nuclear factor 1 (HNF1), which has been shown to bind this site in the HBV large surface antigen promoter (Courtois *et al.*, 1988). HNF1 is a transcription factor which is found predominantly in the liver but is also expressed in the kidney and intestines (Frain *et al.*, 1989; Baumhueter *et al.*, 1990; Kuo *et al.*, 1990) and is known to interact with several liver-specific genes such as albumin, β-fibrinogen and α_1-antitrypsin (Lichtsteiner & Schibler, 1989; Maire *et al.*, 1989; Sawadaishi *et al.*, 1988; Courtois *et al.*, 1987; Courtois *et al.*, 1988; Monaci *et al.*, 1988). In contrast to the effect of the deletion of the HNF1 binding site on the activity of the large surface antigen promoter in the differentiated hepatoma cell lines, the activity of the large surface antigen promoter in the dedifferentiated hepatoma cell line HepG2.1 was not affected by the deletion of the HNF1 recognition sequence. These results are consistent with a similar analysis of the large surface antigen promoter in which the activity of the promoter decreased approximately 30-fold upon deletion of the HNF1 binding site when transfected into the differentiated hepatoma cell lines Huh7 and HepG2, but decreased only 1.4-fold in the poorly differentiated human hepatoma cell line HA22T/VGH (Chang *et al.*, 1989). An effect on the promoter activity was observed in the HepG2.1 cell line only when an additional 34 nucleotides were deleted, leaving only the HBV sequences from -41 to +35 in the construct. This deletion resulted in the loss of 86% of the activity compared with the full-length promoter construct, suggesting that a binding site for a positive acting transcription factor which contributes to the low level of activity in dedifferentiated hepatoma cells may exist between nucleotides -75 and -42. Complete loss of promoter activity was observed in all cell lines when sequences upstream of -25 were deleted. The region between -41 and -25 includes the sequence TATATTATATA, which comprises almost all of the TATATTATATAA sequence defined as the HBV large surface antigen TATA

box element (Fig. VII. 7).

The results of the transient transfection experiments using the four HBV promoters in several cell lines indicated that the nucleocapsid promoter, as well as the large surface antigen promoter, was subject to regulation by liver specific transcription factors (Table VII. 2). Although the activity of the major surface antigen promoter did not appear to be cell line specific in these experiments, a transgenic mouse model has demonstrated liver-specific regulation of the major surface antigen promoter (Araki *et al.*, 1989; Farza *et al.*, 1988). This observation may indicate some of the limitations of the transient transfection system used in this study.

It has been reported that a binding site for HNF1 was located in the enhancer I region, suggesting that HNF1 might contribute to liver cell type specificity of the other HBV promoters (Patel *et al.*, 1989). The evidence for the HNF1 binding site was a DNase 1-protected region, a footprint, using rat liver nuclear extracts on the sequence TCCATTTACACAATGTGGATATCCT (map position 1025 to 1049) just upstream of the classically defined enhancer I (map positions 1070 to 1374). The footprint was competed by an unidentified HNF1 oligonucleotide. The homology to the HNF1 consensus sequence (GTTAATNATTAAC) is not obvious, and the best alignment results in a homology of 5 of 12 conserved nucleotides. It was, however, of interest to determine whether the apparent hepatoma cell line specificity of the nucleocapsid promoter observed in these transient transfection experiments, and the liver specific expression of the major surface antigen (2.1 kb) and nucleocapsid (3.5 kb) RNAs in transgenic mice (Araki *et al.*, 1989; Farza *et al.*, 1988) could be due to the influence of HNF1. To examine the possibility that HNF1 influenced the activity from the other HBV promoters in addition to the large surface antigen promoter, an expression vector containing the cDNA of the rat HNF1 polypeptide was cotransfected with the HBV promoter-LUC constructs into the dedifferentiated hepatoma cell line HepG2.1, which appears to lack the endogenous HNF1 necessary to transactivate the large surface antigen promoter, and into the differentiated hepatoma cell line Huh7, which appears to express functional HNF1 polypeptide. The data demonstrated that the exogenous expression of

HNF1 transactivated only the large surface antigen promoter in both cell lines (Table VII. 3). The HBV major surface antigen, X gene, and nucleocapsid promoters were completely unresponsive to the expression of HNF1 in both cell lines in this system. This is consistent with the observation that none of these three promoters showed dependence upon the HNF1 binding site in the deletion analyses of their transcriptional activities in differentiated hepatoma cell lines (Zhang *et al.*, 1992; Raney *et al.*, 1989; Raney *et al.*, 1991b). The liver cell type specificity of HBV cannot be attributed to the activity of HNF1 alone. Although HNF1 appears to be the factor most involved in the regulation of the large surface antigen promoter, its expression is not sufficient to transactivate the nucleocapsid and surface antigen promoters in this tissue culture system. It is possible that other transcription factors may interact with HNF1 to confer the apparent liver cell type specificity to the nucleocapsid and surface antigen promoters, or that HNF1 is not involved in the regulation of these promoters. Evidence that other transcription factors, including NF-1, EF-C, AP-1, and the liver-enriched C/EBP (Ben-Levy *et al.*, 1989; Dikstein *et al.*, 1990; López-Cabrera *et al.*, 1990; Trujillo *et al.*, 1991; Shaul & Ben Levy, 1987), bind to the enhancer and promoter regions suggests that C/EBP and possibly additional unidentified factors may contribute to the liver-specific regulation of these promoters (Yuh & Ting, 1993; Shaul *et al.*, 1985; Patel *et al.*, 1989). No evidence currently exists to suggest that HNF1 is involved in the transcriptional regulation of the nucleocapsid and surface antigen promoters.

The HNF1 recognition sequence GTTAATCATTACT located from nucleotide -89 to nucleotide -77 upstream of the start of transcription appears to be crucial for the majority of the activity of the large surface antigen promoter in differentiated hepatoma cell lines (Fig. VII. 6). This binding site is also necessary for the induction of a large surface antigen minimal promoter by exogenously expressed HNF1 polypeptide in dedifferentiated and differentiated hepatoma cell lines (Fig. VII. 8). This sequence is completely conserved in all of the sequenced HBV genomes (Fig. V. 3), suggesting it plays an important role in the life cycle of the virus. The lack of dependence on the HNF1 binding site in HepG2.1 cells for PS(1) promoter activity and the relatively lower

PS(1) promoter activity observed in dedifferentiated versus differentiated hepatoma cell lines suggested that the HepG2.1 cells did not express a functional HNF1 polypeptide necessary to increase transcription from the large surface antigen promoter. This is consistent with the dedifferentiated characteristics of the cell line and with a similar observation in which the level of fibrinogen β-chain gene expression was five-fold greater in the rat differentiated hepatoma cell line, Fao, than in the rat dedifferentiated hepatoma cell line, Faof1C2 (Baumhueter *et al.*, 1988). The dedifferentiated variant of the Fao cell line, Faof1C2, expresses a nuclear protein (vHNF1) of lower molecular weight than HNF1 which has DNA binding sequence specificity similar to that of HNF1. The differentiation states of these cell lines correlated with the β-fibrinogen promoter activity, the morphology of the cell lines, and the expression of either the normal (HNF1) or variant (vHNF1) forms of the HNF1 transcription factor (Baumhueter *et al.*, 1988). The liver cell type activity of the large surface antigen promoter appears to be dependent upon the transcription factor HNF1, and may contribute to the hepatotropism of the virus as the large surface antigen is an essential component of the envelope of the viral particle (Ueda *et al.*, 1991). Consistent with this possibility, viral particles have been produced in transfection experiments only in highly differentiated hepatoma cell lines (Sureau *et al.*, 1986; Tsurimoto *et al.*, 1987; Sells *et al.*, 1987; Chang *et al.*, 1987; Yaginuma *et al.*, 1987), which appear to express the HNF1 polypeptide necessary to activate the large surface antigen promoter. The existence of the variant HNF1, or vHNF1, in dedifferentiated hepatoma cell lines, and its ability to bind the HNF1 recognition sequence (Cereghini *et al.*, 1988; Baumhueter *et al.*, 1988) poses questions as to its possible role in the regulation of promoters containing an HNF1 binding site.

Exogenous expression of the HNF1 cDNA in the transient transfection assays confirmed that the HNF1 polypeptide was able to increase the activity of the PS(1) promoter through the HNF1 binding site (Fig. VII. 8). Gel mobility shift analysis was performed to provide physical evidence that the exogenously expressed HNF1 polypeptide could bind to the PS(1) promoter HNF1 recognition sequence, and to determine whether proteins present in the cell lines used for the functional studies could

also bind the PS(1) promoter HNF1 recognition sequence. The results demonstrated that a protein present in Huh7 cells bound the PS(1) promoter HNF1 binding site with similar migration properties to the complex formed between the HNF1 binding site and the exogenously expressed HNF1 polypeptide in the transfected HepG2.1 cell extract (Fig. VII. 9 and 16). These results are consistent with the existence of a functional HNF1 polypeptide in differentiated hepatoma cell lines which can activate transcription through the large surface antigen promoter HNF1 recognition sequence. The complex formed with the Huh7 nuclear extract appeared to comprise a closely migrating doublet. It is possible that the lower band of the doublet represents binding of the variant HNF1 polypeptide, vHNF1, to the HNF1 recognition sequence oligonucleotide. It has been shown that vHNF1 is a slightly smaller polypeptide than HNF1 (apparent molecular weight of 68 to 72 kD compared with 88 to 92 kD) (Baumhueter *et al.*, 1988; Bach *et al.*, 1991; Courtois *et al.*, 1988) which binds the HNF1 recognition sequence and is expressed, at levels much lower than those of HNF1, in liver and differentiated hepatoma cells (Bach *et al.*, 1991; Rey-Campos *et al.*, 1991). It is also possible that the multiple bands in the Huh7 extract represent proteolytic cleavage products or different modified forms of the polypeptide (Chouard *et al.*, 1990; Lichtsteiner & Schibler, 1989). Nuclear extracts prepared from the dedifferentiated HepG2.1 cells and nonhepatoma cell line HeLa formed a complex with the HNF1 binding site oligonucleotide that migrated at the position of the lower band in the doublet. These dedifferentiated hepatoma and nonhepatoma cell lines may express the variant HNF1 polypeptide, vHNF1, which binds the HNF1 recognition sequence and demonstrates slightly faster mobility than HNF1 in gel retardation analyses (Cereghini *et al.*, 1988; Baumhueter *et al.*, 1988). Expression of vHNF1, but not HNF1, by the dedifferentiated hepatoma cell line HepG2.1 is consistent with the observation that dedifferentiated rat hepatoma cell lines, H5 and FaoflC2, express vHNF1 but not HNF1 and are also unable to support transcription from HNF1-dependent promoters (Rey-Campos *et al.*, 1991; Cereghini *et al.*, 1988; Baumhueter *et al.*, 1988). In addition, the complex formed with the HepG2.1 nuclear extract was efficiently competed by addition of 20-fold molar excess of unlabelled HNF1 binding

site oligonucleotide to the gel mobility shift assay (Fig. VII. 9). The complex formed with the HeLa nuclear extract was not competed until 2,000-fold excess oligonucleotide was added, suggesting that the binding of the polypeptide in this extract has a lower specificity for the HNF1 binding site and may not be vHNF1. This would be consistent with the observation of another group that the albumin promoter proximal region containing the HNF1 binding site did not form a gel retardation product with HeLa cell nuclear extract that migrated in the position of either the HNF1 or vHNF1 complexes (Cereghini *et al.*, 1988). These gel mobility shift results indicated that the differentiated hepatoma cell lines, which exhibit HNF1-dependent large surface antigen promoter activity, expressed a polypeptide that could bind the HNF1 recognition sequence and comigrated with the complex produced by the expression of the HNF1 cDNA. These results also demonstrated that the dedifferentiated hepatoma cell line HepG2.1 and the nonhepatoma cell line HeLa did not express the HNF1 polypeptide or related polypeptide with similar migration properties. The HepG2.1 cell line did express, at relatively lower levels, a polypeptide with slightly faster mobility which specifically bound the HNF1 recognition sequence, possibly vHNF1. The HeLa cells appeared to express a polypeptide with similar migration properties to the HepG2.1 polypeptide, but which did not bind to the HNF1 recognition sequence with the same level of specificity.

It has been shown by DNase I footprinting analysis that purified HNF1 and proteins present in differentiated hepatoma cell lines can bind the HBV PS(1) promoter HNF1 recognition sequence and HNF1 binding sites present in other promoters (Courtois *et al.*, 1988; Lichtsteiner & Schibler, 1989; Frain *et al.*, 1989; Baumhueter *et al.*, 1988; Zhou & Yen, 1991). The results of this DNase I footprinting analysis also demonstrated that a protein present in Huh7 nuclear extracts was capable of binding the PS(1) promoter HNF1 recognition sequence (Fig. VII. 10). The Huh7 nuclear extract protected a 30-nucleotide region, spanning nucleotides -99 to -69 on the minus strand, of the PS(1) promoter fragment used in the footprinting assay. The protected region encompassed the 13-nucleotide HNF1 binding site (AGTTTGGA*AGTAATGATTAAC*TAGATGTTCT; the HNF1 site is indicated by italics).

Neither the HepG2.1 nor the HeLa nuclear extracts protected any region of the 423-base-pair PS(1) promoter fragment (Fig. VII. 10). Both of these extracts formed a complex with the HNF1 binding site oligonucleotide in the gel mobility shift assay (Fig. VII. 9), but the complexes migrated in a slightly different position from that of the Huh7 nuclear extract and appeared to be present in much lower amounts than the Huh7 complex. Whether the HepG2.1 and HeLa extracts contain vHNF1 or a different polypeptide that binds the HNF1 recognition sequence, the absence of a footprint using these extracts is probably due to the relatively low level of the polypeptide in these cell lines. The DNase I footprinting analysis also identified the location of the TBP binding site in the HBV PS(1) promoter. Purified TBP (TFIID, Promega) protected a region of the PS(1) promoter from -40 to -25 (*TTATATA*ATATACCCG on the minus strand; the TBP site is indicated by italics), which includes the TATATAA sequence located from -31 to -25 which is similar to eukaryotic TATA box elements.

The protection of the HNF1 recognition sequence by Huh7 nuclear extract, the formation of a gel retardation complex with similar migration properties to the recombinant HNF1, and the dependence on the HNF1 binding site for PS(1) promoter activity in differentiated hepatoma cells are consistent with the suggestion that these differentiated hepatoma cells contain a level of functional HNF1 polypeptide which is sufficient to transactivate the large surface antigen promoter. The absence of a detectable footprint using HepG2.1 and HeLa cell nuclear extracts in the DNase I protection analysis was consistent with the suggestion that the HepG2.1 and HeLa cell lines do not express a level of HNF1 or HNF1-related polypeptide necessary to transactivate the large surface antigen promoter.

The transcription factor HNF1 binds its palindromic target sequence, GTTAATNATTAAC, as a dimer (Frain *et al.*, 1989; Nicosia *et al.*, 1990; Chouard *et al.*, 1990). One might expect that binding of HNF1 to its recognition sequence would be orientation-independent because of these structural features. However, none of the identified HNF1 binding sites conforms to the perfect palindrome of the consensus recognition sequence. Therefore, the influence of the orientation of the HBV large

surface antigen promoter HNF1 binding site on the activity of a minimal promoter element was examined in the absence and presence of exogenously expressed HNF1 in HepG2.1 cells. A double stranded oligonucleotide containing the PS(1) promoter HNF1 binding site was inserted upstream of a PS(1) minimal promoter element, PS(1)pΔ2840-2767LUC, in the same or opposite direction relative to its orientation in the HBV genome (Fig. VII. 11). These constructs were cotransfected with the expression vector containing the HNF1 cDNA, pMTHNF1, or the negative control expression vector, pMT. Comparable levels of transactivation by the HNF1 polypeptide were mediated by the minimal promoters regardless of the orientation of the HNF1 binding site. The full-length and minimal promoter constructs containing the HNF1 binding site were induced seven- to 14-fold by the expression of HNF1, whereas the minimal promoter construct which did not contain the HNF1 binding site, PS(1)pΔ2840-2767LUC, was not induced by the expression of HNF1. These results are consistent with those of two groups who investigated the effect of orientation of the HNF1 binding site using *in vitro* transcription systems. The HNF1 binding sites from the *Xenopus* albumin promoter and the human α1-antitrypsin promoter were both able to support *in vitro* transcription, regardless of their orientation, in the context of a minimal promoter using liver nuclear extracts (Ryffel *et al.*, 1989; Monaci *et al.*, 1988). In this transfection system, the minimal promoter constructs containing the HBV HNF1 binding site demonstrated a relative activity of approximately 10-fold greater than the full-length PS(1) promoter construct and approximately 20-fold greater than the construct deleting through -91 (PS(1)pΔ2840-2718LUC) in the absence of exogenously expressed HNF1 in HepG2.1 cells. This observation suggested that either the position of the HNF1 binding site, perhaps relative to the TATA box element, or the deletion of sequences between -76 and -42, affected the basal activity from this minimal promoter. Although the activity of the PS(1) promoter was not dependent upon the HNF1 recognition sequence in HepG2.1 cells, it is possible that the closer positioning of the binding site to the TATA box element may enable interaction between transcription factors that was not possible at the greater distance, resulting in higher transcriptional activity. It is also possible that a

negative element exists between -76 and -42, which when deleted results in a higher level of basal transcription, but does not affect the inducibility by HNF1. The questions regarding spacing of the HNF1 site and the TATA box and the possibility of a negative element between -76 and -42 were addressed by experiments described in sections VII. M and VII. N. An additional possibility to explain the higher basal activity from these synthetic minimal promoter constructs is that the sequences created at the junction of the oligonucleotide and the minimal promoter construct formed an artificial positive element. The possibility of creating artificial elements always exists when making synthetic promoter constructs and demonstrates the importance of trying to maintain the integrity of the promoter being characterized and to confirm and support the data by performing other experiments.

The ability of the human albumin HNF1 binding site to mediate HNF1 responsiveness was also examined in the context of a PS(1) minimal promoter construct in HepG2.1 cells. A double-stranded oligonucleotide containing the albumin recognition sequence was inserted upstream of the large surface antigen promoter sequences in the construct PS(1)pΔ2840-2767LUC to generate the clone PS(1)pΔ2840-2767AHNF1(+)LUC (Fig. VII. 11). Like the HBV HNF1 recognition sequence, the human albumin HNF1 recognition sequence appears to mediate HNF1 induction by exogenously expressed HNF1 in the context of the HBV PS(1) minimal promoter containing the TATA box element. The level of transcriptional activation was similar to that of the full-length PS(1) promoter construct and of the two constructs containing the HBV HNF1 site inserted upstream of the PS(1) minimal promoter element, indicating that the two elements were comparable in their ability to mediate the influence of HNF1 in this system. The sequences of the HBV PS(1) promoter (GTTAATCATTACT) and human albumin promoter (GTTAATAATCTAC) HNF1 binding sites differ by four of the 12 conserved sequences in the consensus binding site (GTTAATNATTAAC) (Courtois *et al.*, 1988), although each one only differs from the consensus by two nucleotides in the 3' half of the binding site. It appears that these sequence differences do not influence the ability of the transcription factor HNF1 to mediate its effects through these binding

sites. This is not surprising as more than 25 HNF1 binding sites in various liver genes have been identified and none has exactly the same sequence (Tronche & Yaniv, 1992).

The deletion analysis and the transactivation of the synthetic minimal promoter constructs containing the HNF1 recognition sequence demonstrated the importance of the HNF1 element for the activity of the large surface antigen promoter (Fig. VII. 6, 8, 11). The dependence on HNF1 for the liver specific transcriptional activity of several other promoters has been shown in both transient transfection systems and in *in vitro* transcription systems (Cereghini *et al.*, 1988; Cereghini *et al.*, 1990; Courtois *et al.*, 1988; Courtois *et al.*, 1987; Feuerman *et al.*, 1989; Hardon *et al.*, 1988; Ryffel *et al.*, 1989; Tsutsumi *et al.*, 1989; Vaulont *et al.*, 1989). The HNF1 binding site plus a TATA box element has been shown to support *in vitro* transcription using liver nuclear extracts (Ryffel *et al.*, 1989). To determine whether the HBV large surface antigen promoter HNF1 binding site and TATA box element were enough to support transcriptional activity in this transient transfection system, synthetic promoter constructs were made containing the HBV HNF1 recognition sequence, the HBV PS(1) promoter TATA box element, or both, upstream of the luciferase ORF. These plasmids were transfected into HepG2.1 cells in the absence or presence of the HNF1 expression vector and their activities reported relative to that of the full-length PS(1) promoter in the absence of exogenously expressed HNF1. These experiments were performed to determine if these two elements could support transcription in the absence of any of the other HBV sequences, or if additional HBV sequences were needed for their function. The results (Fig. VII. 12) demonstrated that the plasmids containing either the TATA box element or the HNF1 binding site could support transcription, to levels approximately 20% and 40% of the PS(1) promoter, respectively, in the absence of exogenously expressed HNF1. These individual promoter elements can apparently recruit the basal transcriptional machinery effectively enough to generate a low level of activity. The combination of the HNF1 and TATA box elements results in transcriptional activity equivalent to that of the full-length PS(1) promoter in the absence or presence of HNF1. Both of the constructs containing the HNF1 binding site were transactivated to levels similar to the

full-length PS(1) promoter by the expression of HNF1, whereas the construct containing only the TATA box element upstream of the luciferase ORF was not. These results indicate that the HBV PS(1) promoter HNF1 site can mediate transcriptional activation by HNF1 in the absence of all other HBV sequences. The TATA box and HNF1 elements appear to comprise all of the sequences necessary for the regulated transcription from the large surface antigen promoter in this transfection system, and although other factors may exert small modulatory effects which are difficult to detect in this system, it is clear that the transcription factor HNF1 is crucial for the regulation of the large surface antigen gene. These results indicating that HBV sequences other than the HNF1 and TATA box elements are not necessary for HNF1-dependent activity of the PS(1) promoter contradict those which suggest that an Oct-1 transcription factor binding site located 15 nucleotides downstream of the HNF1 binding site is necessary for liver specific transcriptional activity in HepG2 and Huh7 cells and for HNF1 transactivation of the large surface antigen promoter in HeLa cells (Zhou & Yen, 1991). One of the differences in the two analyses which could account for the apparent contradiction is the subtype of the HBV sequence used. In this analysis the subtype *ayw* is used, whereas *adw* was used for the other report. The consensus Oct-1 binding sequence is ATTTGCAT, the *adw* sequence is ATTTACAT, and the *ayw* sequence is ATTTACAC. The Oct-1 binding sequences only differ between these two subtypes by one nucleotide, but that nucleotide decreases the sequence homology of the *ayw* subtype to six of eight nucleotides, whereas the *adw* subtype contains seven of eight of the consensus sequence nucleotides. Additional experiments to address this issue were performed and are described in section VII. N.

The demonstration that the synthetic construct containing the HNF1 recognition sequence and the TATA box element could mediate HNF1-dependent transcriptional activation permitted the analysis of a series of plasmids in which the distance between these two promoter elements was altered by deletion or by addition of linker sequences. The results shown in Fig. VII. 11 raised the question as to whether the spacing between these two binding sites influenced the 10-fold increase in the basal level of activity from

the constructs PS(1)pΔ2840-2767HNF1(+)LUC and PS(1)pΔ2840-2767HNF1(-)LUC. The distance between the HNF1 recognition sequence and the TATA box element in the PS(1) promoter is 45 nucleotides, and in the two constructs mentioned above, it is 22 and 21 nucleotides, respectively. In the synthetic promoter spacing series of constructs, the distance between the HNF1 and TATA elements varied between 10 and 66 nucleotides. The transcriptional activities of the constructs were measured in HepG2.1 cells in the absence and presence of exogenously expressed HNF1 and were compared with the activity of the full-length PS(1) promoter in the absence of exogenously expressed HNF1. The results (Fig. VII. 13) demonstrated that some variation in the levels of activity existed between constructs, and this variation suggests that the distance between the two elements may play a role in determining the level of transcription from these synthetic promoters. As the length of one turn of the DNA helix is approximately 10 nucleotides, the 10-nucleotide periodicity of higher relative activities from the constructs suggested that binding by HNF1 and TBP on the same face of the DNA helix may be preferential for the increased transcriptional activity mediated by these factors. The level of activity from the synthetic construct containing the two elements separated by 45 nucleotides, the distance in the PS(1) promoter, was similar to the activity from PS(1)pLUC (data not shown), suggesting that these two elements, HNF1 and TBP, are sufficient to generate the transcriptional activity of the PS(1) promoter. In addition, transcription from the synthetic promoters did not substantially increase or decrease as the distance changed, even when the distance between the two elements was reduced to only 10 nucleotides, although it appeared to show a general trend towards lower activity as the distance between HNF1 and TBP increased. This observation suggested that as the distance between the HNF1 and TATA box elements increased, the interaction of the factors with the DNA binding elements was less efficient at activating transcription.

Clustered point mutations were made in the region of the large surface antigen promoter between the HNF1 binding site and the TATA box element, between nucleotides -73 and -34. The 5' deletion analysis demonstrated that the sequences upstream of -90 were not necessary for the majority of the transcriptional activity

observed in differentiated hepatoma cell lines, nor were they required for activation by HNF1. Although the spacing constructs demonstrated that only the HNF1 recognition sequence and TATA box element were necessary for transcriptional activity and HNF1 inducibility, it was of interest to determine whether the sequences downstream of the HNF1 binding site had an influence on this activity for two reasons. First, the experiments described in Fig. VII. 11 demonstrated that the uninduced level of activity from the minimal promoter constructs containing the HNF1 binding site [PS(1)pΔ2840-2767HNF1(+)LUC and PS(1)pΔ2840-2767HNF1(-)LUC] was 10-fold higher than the full-length promoter in the absence of HNF1, but the level of induction by HNF1 was essentially equivalent to the full-length PS(1) promoter. This suggested the possibility of a negative element in the region between -76 and -42 that affects the uninduced activity from the promoter. Second, Zhou and Yen (1991) suggested that the Oct-1 sequence, located between nucleotides -61 and -54, was necessary for the HNF1-dependent activity of this promoter in differentiated hepatoma cell lines and for HNF1 inducibility by exogenously expressed HNF1 in HeLa cells. Each of the four mutant constructs, PS(1)pM1LUC to PS(1)pM4LUC, contained a 10-nucleotide mutated region between -76 and -42 (Fig. VII. 14A) and were transfected into HepG2.1 cells in the absence or presence of the HNF1 expression vector, pMTHNF1. All of the mutant constructs were made by PCR synthesis and some contain PCR errors. The PS(1)pM1LUC construct is missing a nucleotide within the mutated region. The resulting clone contains the desired mutation but decreases the distance between the HNF1 binding site and TATA box element by one nucleotide. The data from the spacing constructs indicated that the distance between these two elements was not critical for similar activity, so the PS(1)pM1LUC construct was used. A PCR synthesis error was introduced in the PS(1)pM4LUC construct at nucleotide 2753 (-56), substituting an A for a G.

The results demonstrated that none of the mutations affects the uninduced or HNF1-induced activity from this promoter in HepG2.1 cells to a large extent (Fig. VII. 14B). The levels of activity were within 2.5-fold of the full-length promoter in the absence of HNF1, and the levels of transactivation by HNF1 were within the range

observed for the full-length promoter and the parental deletion construct, PS(1)pΔ2840-2707LUC. Mutation of these sequences did not result in much greater transcriptional activity, suggesting that a negative element does not exist in this region and therefore does not answer the question as to why the PS(1)pΔ2840-2767HNF1LUC constructs show a higher level of uninduced activity in HepG2.1 cells. However, the mutations in these constructs did not cover the entire region that was deleted in the PS(1)pΔ2840-2767HNF1LUC constructs, and it is possible that mutating a larger region, or a different region within this 34-nucleotide segment, could define a negative element if it exists. An alternative explanation for the high uninduced level of activity from the PS(1)pΔ2840-2707LUC construct is that an artificial positive element was created by the juxtaposition of the oligonucleotide sequences and the HBV minimal promoter sequences. The mutation in the construct PS(1)pM2LUC changed six of the eight nucleotides of the Oct-1 binding site, which has been reported to be essential for HNF1 transactivation (Zhou & Yen, 1991), and the relative activities from this construct were essentially equivalent to those of the full-length and parental constructs in the absence or presence of HNF1 in HepG2.1 cells. The mutant constructs, the spacing constructs and the minimal promoter containing the HNF1 binding site yielded the same results, that the Oct-1 binding site is not necessary for HNF1 transactivation in HepG2.1 cells. This apparent contradiction to the results of Zhou and Yen (1991) may be a reflection of differences in the two systems being used. In the study by Zhou and Yen (1991), the *adw* subtype of HBV was used, which has a different Oct-1 sequence (ATTTACAT) from the *ayw* sequence (ATTTACAC) used in this study. The mutant constructs used in the two studies also differed. The *adw* sequence was changed to AAGGACAT, whereas the *ayw* sequence was changed to GGTACCCT. In addition, the cell lines used in these studies were different. Zhou and Yen (1991) presented data from transfection experiments using HepG2 and HeLa cells and reported that Huh7 results were similar, but they did not show the Huh7 data. In this study, Huh7 and HepG2.1 cells were used for the transfection experiments. The HepG2 and Huh7 cells both represent differentiated human hepatoma cell lines, but may have differences that could affect

these transfection results. HeLa cells are human, non-hepatoma cells lacking HNF1, and HepG2.1 cells are dedifferentiated human hepatoma cells which appear to lack levels of functional HNF1 necessary for activating transcription from the HBV large surface antigen promoter. It seems possible that these two cell lines might differ enough that the results of transient transfections using these cell lines for the exogenous expression of HNF1 may differ.

All of these mutant constructs were also transfected into Huh7 cells and their activities compared with those of the full-length and PS(1)pΔ2840-2707LUC constructs. As in the HepG2.1 cells, the activities of PS(1)pM3LUC and PS(1)pM4LUC were approximately the same as the control plasmids. The activities of the PS(1)pM1LUC and PS(1)pM2LUC constructs were approximately four- to five-fold lower than the control plasmids in Huh7 cells. It is possible that the proximity to the HNF1 recognition sequence of these mutations (three to 12 and 13 to 22 nucleotides, respectively) affects the efficiency of endogenous HNF1 binding to its recognition sequence and reduces the activity level by approximately four-fold. This effect was not observed in the HepG2.1 cells in the presence of exogenously expressed HNF1, possibly due to a difference in the level of transcription factor present when it is exogenously expressed. The level of exogenously expressed HNF1 in HepG2.1 cells may be higher than the endogenous HNF1 in Huh7 cells, and HNF1 may be present in enough excess that it can overcome a reduced efficiency of binding so that no effect is observed in HepG2.1 cells. These mutated constructs did retain relatively high levels of activity when compared with the constructs lacking the HNF1 binding site. These results are inconsistent with those demonstrating the importance of the Oct-1 binding site in HNF1-dependent activity in HepG2 cells (Zhou & Yen, 1991). In that analysis, mutation of the Oct-1 binding site resulted in a greater than 10-fold decrease in activity in HepG2 cells, to levels similar to the activity in HeLa cells. In this analysis, the relative activity decreased only four-fold upon mutation of the Oct-1 site in Huh7 cells, and this relative activity is 50- to 70-fold higher than the relative activity of the PS(1) promoter in HepG2.1 or HeLa cells (Table VII. 2). These data are consistent with the HepG2.1 data in that the HNF1-dependent

activity from the PS(1) promoter observed in these cell lines is not dependent upon the Oct-1 site, whereas in the study by Zhou and Yen (1991), the Oct-1 site is essential for HNF1-dependent activation of PS(1) promoter transcription.

The importance of the HNF1 binding site to the activity of the HBV large surface antigen promoter has been demonstrated by deletion analysis (Fig. VII. 6 and 8). The construct PS(1)pΔ2840-2425M5LUC, which contains a mutation of the HNF1 recognition sequence, was made to determine whether other HBV sequences could compensate for the activity of the HNF1 binding site in its absence. The results, shown in Fig. VII. 14, were consistent with those of the deletion analyses indicating that the HNF1 binding site was not necessary for the low level of transcriptional activity present in HepG2.1 cells, but was absolutely critical for transactivation of the promoter in the presence of exogenously expressed HNF1. In Huh7 cells, the mutant construct PS(1)pΔ2840-2425M5LUC had essentially no transcriptional activity, which was also consistent with the suggestion that the HNF1 binding site is the mediator of the HNF1-dependent activity of the large surface antigen promoter. These experiments were also performed with a construct which contains the HNF1 (M5) mutation in the context of the full-length genome and the results were the same as with the PS(1)pΔ2840-2425M5LUC plasmid. The results were not shown because the full-length M5 plasmid, PS(1)pM5LUC, contained three point mutations as a result of PCR synthesis. The point mutations were in regions previously shown not to be important for activity of this promoter, but their effect in this plasmid is unknown. It is apparent from these experiments that the HNF1 recognition sequence located between nucleotides -90 and -77, and no other promoter or enhancer element, is the critical component in the HNF1-dependent activity of the large surface antigen promoter.

The cloning of the HNF1 cDNA permitted the analysis of the ability of the HNF1 polypeptide to transactivate promoters containing the HNF1 recognition sequence. The ability to demonstrate transactivation by the exogenously expressed HNF1 polypeptide permitted the characterization of the region of the polypeptide mediating this transcriptional activation. A series of deletion mutants of the HNF1 cDNA in the

expression vector pMTHNF1 was tested for its ability to activate transcription in HepG2.1 cells from plasmids containing the HNF1 recognition sequence (Fig. VII. 15). Gel retardation analysis was performed using nuclear extracts prepared from HepG2.1 cells transfected with the truncated HNF1 expression vectors to demonstrate that the polypeptides being produced were able to bind the HNF1 recognition sequence. The results (Fig. VII. 16) demonstrated that polypeptides produced by the truncated expression vectors formed complexes with the HNF1 binding site that decreased in size corresponding to the extent of the cDNA deletion. The complex formed with the nuclear extract containing the full-length polypeptide comigrated with the band observed in untransfected Huh7 cells, suggesting that the Huh7 cells express HNF1. These experiments confirmed that the expression vectors that did not show transactivation did produce polypeptides capable of binding the HNF1 recognition sequence. The results of the gel mobility shift analysis suggested that transfection of the HNF1 expression vectors may result in the synthesis of different amounts the truncated polypeptides. To determine whether the amount of expression vector transfected affected the activation of the large surface antigen promoter, increasing amounts of HNF1 expression vectors were transfected into HepG2.1 cells. The results of this experiment demonstrated (Fig. VII. 17) that the level of induction by HNF1 increased as the amount of transfected vector increased and that increasing the amount of vectors which did not transactivate did not affect the PS(1) promoter activity. The mapping of the transactivation domains of the HNF1 polypeptide was not affected by the amount of expression vector transfected. The gel retardation complex formed with untransfected HepG2.1 nuclear extract also demonstrated a low level expression of a specific DNA binding polypeptide with slightly faster migration properties than HNF1, suggesting the presence of vHNF1 in the dedifferentiated HepG2.1 cells. This is consistent with the expression pattern of vHNF1 in other dedifferentiated hepatoma cell lines (Cereghini *et al.*, 1988; Mendel & Crabtree, 1991; Baumhueter *et al.*, 1988; De Simone *et al.*, 1991).

Several structural features of the HNF1 polypeptide have been characterized and the HNF1 cDNA deletion constructs used to identify the activation domain were

designed to retain the elements of the polypeptide in the amino-terminal half of the polypeptide which were required for sequence-specific DNA binding. These elements include a dimerization domain in the amino-terminal 32 amino acid residues, a region with homology to the A subregion of the POU-specific domain present in Pit1/GHF1, Oct-1 and Oct-2, and Unc-86, and a divergent 81-amino acid homeodomain in the amino-terminal 281 amino acids (Frain *et al.*, 1989; Chouard *et al.*, 1990; Nicosia *et al.*, 1990). A transcriptional activation domain was identified between amino acids 393 and 517, with the region between amino acids 458 and 476 contributing the majority of the transactivation of the large surface antigen promoter. Recently, a transactivation domain corresponding to the one identified in this study was mapped in HNF1 cotransfection experiments in HeLa cells using the albumin and C-reactive protein HNF1 binding sites (Toniatti *et al.*, 1993). The critical region between amino acids 458 and 476 contains a high concentration of glutamine and proline residues. These two amino acids comprise 50% of the residues in this 18-amino acid region. Several transcription factors appear to have activation domains characterized by high concentrations of particular amino acids, such as the proline-rich domains of AP-2 and CTF/NF-1, the glutamine-rich activation domains of Sp1, Oct-1, and Oct-2A, and the negatively charged domains of AP-1, VP16 and GAL4 (Hope & Struhl, 1986; Ma & Ptashne, 1987; Triezenberg *et al.*, 1988a; Williams *et al.*, 1988; Seipel *et al.*, 1992; Mermod *et al.*, 1989; Courey & Tjian, 1988; Mitchell & Tjian, 1989; Kadonaga *et al.*, 1988; Williams & Tjian, 1991; Bohmann & Tjian, 1989). It is thought that the activation domains function by mediating interactions between the sequence specific transcription factors and other proteins, either components of the general transcription machinery, or coactivators. Different types of activation domains may interact with different components of the general transcription machinery or may communicate with specific coactivators, which may increase the rate of transcription initiation. The Sp1 glutamine-rich activation domain appears to interact with a TBP:TAF_{II}250:TAF_{II}110 complex in the *Drosophila* system (Weinzierl *et al.*, 1993). Recent evidence supports a model for different functional mechanisms of action for different classes of activation domain (Seipel *et al.*, 1992). It appears that glutamine-

rich activation domains act from sites close to the promoter but do not work from remote positions, as an enhancer element might. Negatively charged domains, such as those in VP16, GAL4, and p65 (NF-kβ), and the serine/threonine-rich domains of ITF-1 and ITF-2 were able to activate transcription from remote (i.e., enhancer) as well as proximal positions, and proline-rich domains typical of CTF/NF1 and AP-2 were good proximal activators and weak activators from remote positions (Seipel *et al.*, 1992). Analyses such as this may help to characterize the modes of action used by different transcription factors based on the structure of their activation domains. These modular activation domains may interact with different coactivators/TAFs or strengthen interactions with targets present in limiting concentrations. It has been suggested that the factors containing "proximal" activation domains may interact directly with the basal transcription machinery and that different classes of activation domains, such as the acidic, glutamine-rich and proline-rich domains, interact with different proteins in the preinitiation complex. Based on the identification of the activation domain of HNF1 in this analysis, it would appear that the HNF1 activation domain might fall into the "proximal" position activator group, unable to activate from a remote position unless the glutamine/proline activation domain of HNF1 possesses the weak distal activation characteristics of the proline-rich domains identified. In HBV, HNF1 appears to act from a proximal position. The HNF1 protein bound to the HNF1 recognition sequence located at -89 to -78 upstream of the large surface antigen transcription initiation site mediates strong transcriptional transactivation, whereas the same site which is located 427, 1760, and 2255 nucleotides upstream of the major surface antigen, X gene, and nucleocapsid transcriptional start sites does not affect transcription from these promoters (Raney *et al.*, 1991a).

The results of this analysis differed from a previous study which identified the transactivation domains of the HNF1 polypeptide using an *in vitro* transcription system with the albumin promoter HNF1 recognition sequence (Nicosia *et al.*, 1990). The regions identified as important for transcriptional activation in the *in vitro* system were located from amino acids 281 to 318 and from amino acids 547 to 628. These regions

had no effect on the activation by the HNF1 polypeptide of the three promoter constructs analyzed in this transient transfection system, including a minimal promoter containing the albumin HNF1 binding site. The region between amino acids 502 and 516 had a small effect, less than two-fold, on the activation level mediated by HNF1 in the transfection system, but no effect was observed upon deletion of amino acids 548 to 628. The results of the internal deletions (Fig. VII. 15) confirmed those of the 3' deletions, demonstrating that neither the region from residues 281 to 318 nor the region from 547 to 628 mediated the activation from the HNF1 polypeptide in the transfection system, thus defining mutually exclusive regions of the polypeptide as activation domains (Tronche & Yaniv, 1992). This result is interesting because the mapping of the Sp1 activation domains concurred in *in vitro* transcription and transient transfection systems (Courey & Tjian, 1988; Kadonaga *et al.*, 1988). The difference in the sequences of the HNF1 binding site cannot account for the different results because the albumin HNF1 binding site was used in both analyses. A possible explanation which could account for the apparent differences in the characterization of the activation domains of HNF1 is the differences in the nature of the systems used. In the *in vitro* transcription study, a rat spleen nuclear extract which was complemented by nuclear extract from HNF1-recombinant vaccinia virus infected HeLa cells, was supplied with a template containing polymerized HNF1 binding sites from the albumin promoter inserted next to the herpes simplex virus thymidine kinase (HSV-TK) promoter TATA box element. In this study, the dedifferentiated hepatoma cell line HepG2.1 was cotransfected with the promoter construct and the expression vector containing the HNF1 cDNA. It is likely that the composition of transcription factors in the spleen nuclear extract complemented with vaccinia virus HeLa cell nuclear extract is different from the HepG2.1 cells in culture. Different cell-type specific coactivators present in each system may be mediating the activity of HNF1 by interacting with distinct regions of the polypeptide, resulting in the mapping of different domains which are important for transactivation in the two systems. Evidence that a cell-type-specific inhibitor is involved in the control of c-*Jun* activity has resolved a similar discrepancy in the findings of a transfection

study using F9 teratocarcinoma cells (Angel *et al.*, 1989) and an *in vitro* study using HeLa cell extracts (Bohmann & Tjian, 1989) mapping the activation domains of c-*Jun*. It was observed that an activation domain, A1, of c-*Jun* was necessary for transactivation of AP-1 containing promoter elements in transfection of F9 cells (Angel *et al.*, 1989), whereas this region had no activity in the HeLa *in vitro* system (Bohmann & Tjian, 1989). Further experiments have demonstrated that the A1 site is an activation domain and increases activation in *Drosophila* Schneider line 2 (SL2) cells, human hepatoma HepG2 cells and F9 cells in transfection experiments (Baichwal & Tjian, 1990). This region also activated transcription in an *in vitro* transcription system using *Drosophila* SL2 extracts but not when using HeLa extracts (Baichwal & Tjian, 1990). Competition experiments demonstrated that a cell-type-specific inhibitor in HeLa cells represses the activity of the A1 activation domain and explains the differences in the initial mapping of the activation domains of c-*Jun* (Angel *et al.*, 1989; Bohmann & Tjian, 1989). The inhibitor appears to interact with the negative regulatory domain, δ, and an additional domain (ϵ), which are adjacent to the A1 domain, to modulate the activity of A1 (Baichwal *et al.*, 1992). This inhibitory activity is not present in the other cell types examined (Angel *et al.*, 1989; Baichwal & Tjian, 1990; Baichwal *et al.*, 1992). Very recently, a transfection study was reported (Toniatti *et al.*, 1993) which confirms the presence of an HNF1 activation domain in the glutamine/proline-rich region of the HNF1 polypeptide identified as critical for mediating transactivation of the HBV large surface antigen promoter (Fig. VII. 15).

A transcription factor related to HNF1, variant HNF1 (vHNF1), binds the same recognition sequence as HNF1, can transactivate a promoter containing the HNF1 binding site, and can form heterodimers with HNF1 (Baumhueter *et al.*, 1988; De Simone *et al.*, 1991; Rey-Campos *et al.*, 1991; Mendel *et al.*, 1991a). vHNF1 is expressed in differentiated and dedifferentiated hepatoma cell lines and in the same tissues as HNF1 such as liver, kidney and intestine, although at different levels (Bach *et al.*, 1991; Rey-Campos *et al.*, 1991). The cDNA for vHNF1 has been cloned and sequenced and the predicted protein exhibits structural features characteristic of homeoproteins (Rey-

Campos *et al.*, 1991). The highest degree of sequence homology to HNF1 exists in the regions important for dimerization and specific DNA binding. The carboxy-terminal portion of the proteins do not contain a high level of homology, and there is no homology in the regions of the activation domains mapped by *in vitro* transcription (Nicosia *et al.*, 1990; De Simone *et al.*, 1991; Rey-Campos *et al.*, 1991; Mendel *et al.*, 1991a; Bach *et al.*, 1991). As the vHNF1 polypeptide is only 557 amino acids long, it lacks the carboxyl-terminal portion of the polypeptide which contains one of the two previously defined activation domains. Interestingly, one region in the carboxy-terminal half of the polypeptide which shows homology to HNF1 is in the region between residues 393 and 548, the activation domain identified in this study. This region displays 54% amino acid identity (71 of 132 residues), and the region critical for activation of the large surface antigen promoter, between amino acid residues 458 and 475, displays 78% amino acid identity. The conservation of these regions suggests that the transactivation function of the vHNF1 polypeptide could be mediated by the same domain as the one identified in HNF1 on the large surface antigen and albumin HNF1 recognition sequences in transient transfection assays. vHNF1 was initially thought to exhibit a different tissue distribution from HNF1 and the two appeared to be mutually exclusive in differentiated versus dedifferentiated hepatoma cell lines. However, it has recently been shown that vHNF1 is expressed in liver tissue and differentiated hepatoma cell lines (Bach *et al.*, 1991; Rey-Campos *et al.*, 1991), and that vHNF1 can transactivate the albumin promoter in transient cotransfection experiments (Rey-Campos *et al.*, 1991). It would be very interesting to characterize the domain responsible for this activation to determine whether it is the same or a different domain from the one utilized by HNF1 in transient transfection systems. This study has been initiated by inserting the vHNF1 cDNA into an expression vector to perform the deletion analysis of the cDNA in an effort to identify the transcriptional activation domain of vHNF1 in transient transfection experiments.

Heterodimerization has been shown to be an effective mechanism to alter the activity and binding specificity of transcription factors. The AP-1 family of proteins,

including the *Fos* proteins and *Jun* proteins, heterodimerize, and *Fos/Jun* heterodimers bind DNA and activate transcription more efficiently than *Jun* homodimers (Curran & Franza, 1988). The heterodimerization of two homeodomain proteins, the yeast mating type factor Mat a1 and Mat α2, appears to change the binding specificity of Mat α2 (Goutte & Johnson, 1988; Dranginis, 1990). Similarly, the binding specificity of two *Jun* heterodimers, *Jun/Fos* and *Jun*/CREB, is different. The *Jun/Fos* heterodimer preferentially recognizes AP-1 sites over the related CRE recognition sequence (Macgregor *et al.*, 1990). In cotransfection experiments, it appears that heterodimerization of vHNF1 and HNF1 does not have a strong effect on the transactivation of the albumin promoter proximal element (Rey-Campos *et al.*, 1991; De Simone *et al.*, 1991; Mendel *et al.*, 1991a). The relative amounts of vHNF1 and HNF1 homodimers and heterodimers in different tissues may affect the activities of HNF1-binding-site-containing genes in different tissues. Heterodimerization could also broaden the spectrum of coactivators or general transcription factors through which these factors might act.

The results of the studies described here indicate that the transcriptional regulation of the HBV large surface antigen gene appears to be relatively simple. The viral enhancer elements have little, if any, effect on this promoter, and the minimal regulatory region appears to comprise two critical binding elements, the hepatocyte nuclear factor 1 recognition sequence and the TATA box element. Other elements may subtly modulate the activity from this promoter, but in the transfection system used for this study, the majority of transcription from the large surface antigen promoter appears to be mediated by the factors that bind the HNF1 and TATA box elements.

The role of liver-specific, or liver-enriched transcription factors in the regulation of HBV transcription is central to the life cycle of the virus. Generation of the replication template is dependent upon the nucleocapsid promoter, which exhibits liver cell type specificity, and production of the infectious viral particle is dependent upon the assembly of the correct proportion of viral polypeptides. As the large surface antigen is an essential component of the virion, the tissue-specific regulation of its rare

mRNA is likely to be an extremely important aspect of the viral life cycle. HNF1, like its homeoprotein counterparts shown to be important in the development of many organisms, plays a key role in the appropriate regulation of the expression of this mRNA and is therefore a major influential factor in the life cycle of HBV. The regulation of HNF1 itself and the means by which HNF1 mediates its effects, through coactivators or by direct interaction with the general transcription machinery, or by heterodimerization with other factors to alter its specificity, are areas which provide many opportunities for studying both the biology of HBV and the more basic mechanisms of transcriptional regulation.

IX. REFERENCES

ACS, G., SELLS, M.A., PURCELL, R.H., PRICE, P., ENGLE, R., SHAPIRO, M. & POPPER, H. (1987). Hepatitis B virus produced by transfected HepG2 cells causes hepatitis in chimpanzees. *Proceedings of the National Academy of Sciences, U.S.A.* **84**, 4641-4644.

ADEN, D.P., FOGEL, A., PLOTKIN, S., DAMJANOV, I. & KNOWLES, B.B. (1979). Controlled synthesis of HBsAg in a differentiated human liver carcinoma-derived cell line. *Nature, London* **282**, 615-616.

ALBERTI, A., TREVISIAN, A., FATTOVICH, G. & REALDI, G. (1984). The role of hepatitis B virus replication and hepatocyte membrane expression in the pathogenesis of HBV-related hepatic damage. In *Advances in Hepatitis Research*, pp. 134-143. Edited by F. V. Chisari. New York: Masson.

ALBERTI, A., PONTISSO, P., TAGARIELLO, G., CAVALLETTO, D., CHEMELLO, L. & BELUSSI, F. (1988). Antibody response to Pre-S2 and hepatitis B virus induced liver damage. *Lancet* **1**, 1421-1424.

ALBIN, C. & ROBINSON, W.S. (1980). Protein kinase activity in hepatitis B virus. *Journal of Virology* **34**, 297-302.

ALEXANDER, J., BEY, E., WHITCUTT, J.M. & GEAR, J.H.S. (1976). Adaption of cells derived from human malignant tumours to growth in vitro. *South African Journal of Medical Sciences* **41**, 89-98.

ALMEIDA, J.D., RUBENSTEIN, D. & STOTT, E.J. (1971). New antigen-antibody system in Australia-antigen-positive hepatitis. *Lancet* **2**, 1225-1227.

ALMEIDA, J.D. & WATERSON, A.P. (1969). Immune complexes in hepatitis. *Lancet* **2**, 983-986.

ANGEL, P., SMEAL, T., MEEK, J. & KARIN, M. (1989). Jun and v-jun contain multiple regions that participate in transcriptional activation in an interdependent manner. *New Biologist* **1**, 35-43.

ANTONUCCI, T.K. & RUTTER, W.J. (1989). Hepatitis B virus (HBV) promoters are regulated by the HBV enhancer in a tissue-specific manner. *Journal of Virology* 63, 579-583.

ARAKAKI, Y., TOYODA, H., WATANABE, S., SEKI, S. & ODA, T. (1988). Detection of hepatitis B virus-DNA in liver tissues in chronic type B liver diseases and its relevancy to viral antigens: A study by in situ hybridization using a biotinylated probe and streptavidin-alkaline phosphatase. *Acta Histochemistry and Cytochemistry* 21, 489-498.

ARAKI, K., MIYAZAKI, J-I., HINO, O., TOMITA, N., CHISAKA, O., MATSUBARA, K. & YAMAMURA, K-I. (1989). Expression and replication of hepatitis B virus genome in transgenic mice. *Proceedings of the National Academy of Sciences, U.S.A.* 86, 207-211.

ASSA-MUNT, N., MORTISHIRE-SMITH, R.J., AURORA, R., HERR, W. & WRIGHT, P.E. (1993). The solution structure of the Oct-1 POU-specific domain reveals a striking similarity to the bacteriophage lambda repressor DNA-binding domain. *Cell* 73, 193-205.

AURORA, R. & HERR, W. (1992). Segments of the POU domain influence one another's DNA-binding specificity. *Molecular and Cellular Biology* 12, 455-467.

AUSUBEL, F.M., BRENT, R., KINGSTON, R.E., MOORE, D.D., SEIDMAN, J.G., SMITH, J.A. & STRUHL, K. (1987). *Current Protocols in Molecular Biology*. New York: John Wiley and Sons.

BABINET, C., FARZA, H., MORELLO, D., HADCHOUEL, M. & POURCEL, C. (1985). Specific expression of hepatitis B surface antigen (HBsAg) in transgenic mice. *Science* 230, 1160-1163.

BACH, I., MATTEI, M.-G., CEREGHINI, S. & YANIV, M. (1991). Two members of an HNF1 homeoprotein family are expressed in human liver. *Nucleic Acids Research* 19, 3553-3559.

BAICHWAL, V.R., PARK, A. & TJIAN, R. (1992). The cell-type-specific activator region of c-Jun juxtaposes constitutive and negatively regulated domains. *Genes and Development* 6, 1493-1502.

BAICHWAL, V.R. & TJIAN, R. (1990). Control of c-jun activity by interaction of a cell-specific inhibitor with regulatory domain δ: differences between v- and c-jun. *Cell* 63, 815-825.

BANCROFT, W.H., SNITBHAN, R., SCOTT, R.MCN., TINGPALAPONG, M., WATSON, W.T., TANTICHAROENYOS, P., KARWACKI, J.J. & SCRIMARUT, S. (1977). Transmission of hepatitis B virus to gibbons by exposure to human saliva containing hepatitis B surface antigen. *Journal of Infectious Diseases* 135, 79-85.

BARKER, L.F., CHISARI, F.V., MCGRATH, P.P., DALGARD, D.W., KIRSCHSTEIN, R.L., ALMEIDA, J.D., EDGINGTON, T.S., SHARP, D.G. & PETERSON, M.R. (1973). Transmission of type B viral hepatitis to chimpanzees. *Journal of Infectious Diseases* 127, 648-662.

BARKER, L.F., MAYNARD, J.E., PURCELL, R.H., HOOFNAGLE, J.H., BERQUIST, K.R. & LONDON, W.T. (1975). Viral hepatitis, type B, in experimental animals. *American Journal of the Medical Sciences* 270, 189-194.

BARTENSCHLAGER, R. & SCHALLER, H. (1988). The amino-terminal domain of the hepadnaviral P-gene encodes the terminal protein (genome-linked protein) believed to prime reverse transcription. *EMBO Journal* 7, 4185-4192.

BAUMHUETER, S., COURTOIS, G. & CRABTREE, G.R. (1988). A variant nuclear protein in dedifferentiated hepatoma cells binds to the same functional sequences in the β fibrinogen gene promoter as HNF1. *EMBO Journal* 7, 2485-2493.

BAUMHUETER, S., MENDEL, D.B., CONLEY, P.B., KUO, C.J., TURK, C., GRAVES, M.K., EDWARDS, C.A., COURTOIS, G. & CRABTREE, G.R. (1990). HNF-1 shares three sequence motifs with the POU domain proteins and is identical to LF-B1 and APF. *Genes and Development* 4, 372-379.

BAVAND, M., FEITELSON, M. & LAUB, O. (1989). The hepatitis B virus-associated reverse transcriptase is encoded by the viral pol gene. *Journal of Virology* 63, 1019-1021.

BAYER, M.E., BLUMBERG, B.S. & WERNER, B. (1968). Particles associated with Australia antigen in the sera of patients with leukaemia, Down's syndrome and hepatitis. *Nature, London* 218, 1057-1059.

BEASLEY, R.P., HWANG, L-Y., LIN, C-C. & CHIEN, C-S. (1981). Hepatocellular carcinoma and hepatitis B virus - A prospective study of 22707 men in Taiwan. *Lancet* 2, 1129-1133.

BEASLEY, R.P. (1988). Hepatitis B virus: The major etiology of hepatocellular carcinoma. *Cancer* 61, 1942-1956.

BEN-LEVY, R., FAKTOR, O., BERGER, I. & SHAUL, Y. (1989). Cellular factors that interact with the hepatitis B virus enhancer. *Molecular and Cellular Biology* 9, 1804-1809.

BENDER, W., AKAM, M., KARCH, F., BEACHY, P.A., PEIFER, M., SPIERER, P., LEWIS, E.B. & HOGNESS, D.S. (1983). Molecular genetic of the bithorax complex in Drosophila melanogaster. *Science* 221, 23-29.

BERGER, S.L., CRESS, W.D., CRESS, A., TRIEZENBERG, S.J. & GUARENTE, L. (1990). Selective inhibition of activated but not basal transcription by the acidic activation domain of VP16: Evidence for transcriptional adaptors. *Cell* 61, 1199-1208.

BICHKO, V., PUSHKO, P., DREILINA, D., PUMPEN, P. & GREN, E. (1985). Subtype ayw variant of hepatitis B virus: DNA primary structure analysis. *FEBS Letters* 185, 208-212.

BLUM, H.E., STOWRING, L., FIGUS, A., MONTGOMERY, C.K., HAASE, A.T. & VYAS, G.N. (1983). Detection of hepatitis B virus DNA in hepatocytes, bile duct epithelium, and vascular elements by in situ hybridization. *Proceedings of the National Academy of Sciences, U.S.A.* 80, 6685-6688.

BLUM, H.E., HASSE, A.T. & VYAS, G.N. (1984). Molecular pathogenesis of hepatitis B virus infection: simultaneous detection of viral DNA and antigens in paraffin-embedded liver sections. *Lancet* 2, 771-775.

BLUM, H.E., ZHANG, Z.-S., GALUN, E., VON WEIZSÄCKER, F., GARNER, B., LIANG, T.J. & WANDS, J.R. (1992). Hepatitis B virus X protein is not central to the viral life cycle in vitro. *Journal of Virology* 66, 1223-1227.

BLUMBERG, B.S., ALTER, H.J. & VISNICH, S. (1965). A new antigen in leukemia sera. *Journal of the American Medical Association* 191, 541-546.

BLUMBERG, B.S., SUTNICK, A.I. & LONDON, W.T. (1969). Australia antigen and hepatitis. *Journal of the American Medical Association* **207**, 1895-1896.

BLUMENFELD, M., MAURY, M., CHOUARD, T., YANIV, M. & CONDAMINE, H. (1991). Hepatic nuclear factor 1 (HNF1) shows a wider distribution than products of its known target genes in developing mouse. *Development* **113**, 589-599.

BODNER, M., CASTRILLO, J.L., THEILL, L.E., DEERINCK, T., ELLISMAN, M. & KARIN, M. (1988). The pituitary-specific transcription factor GHF-1 is a homeobox-containing protein. *Cell* **55**, 505-518.

BOHMANN, D. & TJIAN, R. (1989). Biochemical analysis of transcriptional activation by Jun: differential activity of c- and v-Jun. *Cell* **59**, 709-717.

BOND, H.E. & HALL, W.T. (1972). Separation and purification of hepatitis-associated antigen into morphologic types by zonal ultracentrifugation. *Journal of Infectious Diseases* **125**, 263-268.

BOSCH, V., BARTENSCHLAGER, R., RADZIWILL, G. & SCHALLER, H. (1988). The duck hepatitis B virus P-Gene codes for protein strongly associated with the 5'-end of the viral DNA minus strand. *Virology* **166**, 475-485.

BRADFORD, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248-254.

BRAHM, J., VENTO, S., RONDANELLI, E.G., RANIERI, S., FAGAN, E.A., WILLIAMS, R. & EDDLESTON, A.L.W.F. (1988). Sequential studies of Pre-S2 antigenemia and anti-Pre-S2 antibodies in relation to viral replication in acute hepatitis B followed from the early incubation phase. *Journal of Medical Virology* **24**, 205-209.

BREATHNACH, R. & CHAMBON, P. (1981). Organization and expression of eucaryotic split genes coding for proteins. *Annual Review of Biochemistry* **50**, 349-383.

BRIGGS, M.R., KADONAGA, J.T., BELL, S.P. & TJIAN, R. (1986). Purification and biochemical characterization of the promoter-specific transcription factor, SP1. *Science* **234**, 47-52.

BRUSS, V. & GERLICH, W.H. (1988). Formation of transmembranous hepatitis B e-antigen by cotranslational in vitro processing of the viral precore protein. *Virology* **163**, 268-275.

BUDKOWSKA, A., SHIH, J.W-K. & GERIN, J.L. (1977). Immunochemistry and polypeptide composition of hepatitis B core antigen (HBcAg). *Journal of Immunology* **118**, 1300-1305.

BUDKOWSKA, A., DUBREUIL, P., OUATARRA, A. & PILLOT, J. (1988). Anti-Pre-S2 as only serum HBV marker: Possible relation to HBV -2 infection. *Lancet* **1**, 656.

BULLA, G.A. & SIDDIQUI, A. (1988). The hepatitis B virus enhancer modulates transcription of the hepatitis B virus surface antigen gene from an internal location. *Journal of Virology* **62**, 1437-1441.

BURATOWSKI, S., HAHN, S., GUARENTE, L. & SHARP, P.A. (1989). Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* **56**, 549-561.

BURRELL, C.J., GOWANS, E.J., JILBERT, A.R., LAKE, J.R. & MARMION, B.P. (1982). Hepatitis B virus DNA detection by In situ cytohybridization: implications for viral replication strategy and pathogenesis of chronic hepatitis. *Hepatology* **2**, 85S-91S.

CARRASCO, A.E., MCGINNIS, W., GEHRING, W.J. & DEROBERTIS, E.M. (1984). Cloning of an X. laevis gen expressed during early embryogenesis coding for a peptide region homologous to Drosophila homeotic genes. *Cell* **37**, 409-414.

CATTANEO, R., WILL, H., HERNANDEZ, N. & SCHALLER, H. (1983). Signals regulating hepatitis B surface antigen transcription. *Nature, London* **305**, 336-338.

CATTANEO, R., WILL, H. & SCHALLER, H. (1984). Hepatitis B virus transcription in the infected liver. *EMBO Journal* **3**, 2191-2196.

CEREGHINI, S., BLUMENFELD, M. & YANIV, M. (1988). A liver-specific factor essential for albumin transcription differs between differentiated and dedifferentiated rat hepatoma cells. *Genes and Development* **2**, 957-974.

CEREGHINI, S., YANIV, M. & CORTESE, R. (1990). Hepatocyte dedifferentiation and extinction is accompanied by a block in the synthesis of mRNA coding for the transcription factor HNF1/LFB1. *EMBO Journal* **9**, 2257-2263.

CEREGHINI, S., OTT, M.-O., POWER, S. & MAURY, M. (1992). Expression patterns of vHNF1 and HNF1 homeoproteins in early postimplantation embryos suggest distinct and sequential developmental roles. *Development* **116**, 783-797.

CHAKRABORTY, P.R., RUIZ-OPAZO, N., SHOUVAL, D. & SHAFRITZ, D.A. (1980). Identification of integrated hepatitis B virus DNA and expression of viral RNA in an HBsAg-producing human hepatocellular carcinoma cell line. *Nature, London* **286**, 531-533.

CHAKRABORTY, P.R., RUIZ-OPAZO, N. & SHAFRITZ, D.A. (1981). Transcription of human hepatitis B virus core antigen gene sequences in an in vitro HeLa cellular extract. *Virology* **111**, 647-652.

CHANG, C., JENG, K-S., HU, C-P., LO, S.J., SU, T-S., TING, L-P., CHOU, C-K., HAN, S-H., PFAFF, E., SALFELD, J. & SCHALLER, H. (1987). Production of hepatitis B virus in vitro by transient expression of cloned HBV DNA in a hepatoma cell line. *EMBO Journal* **6**, 675-680.

CHANG, H-K., WANG, B-Y., YUH, C-H., WEI, C-L. & TING, L-P. (1989). A liver-specific nuclear factor interacts with the promoter region of the large surface protein gene of human hepatitis B virus. *Molecular and Cellular Biology* **9**, 5189-5197.

CHANG, H.K., CHOU, C.K., CHANG, C., SU, T.S., HU, C., YOSHIDA, M. & TING, L.P. (1987). The enhancer sequence of human hepatitis B virus can enhance the activity of its surface gene promoter. *Nucleic Acids Research* **15**, 2261-2268.

CHANG, H.K. & TING, L.P. (1989). The surface gene promoter of the human hepatitis B virus displays a preference for differentiated hepatocytes. *Virology* **170**, 176-183.

CHANG, L-J., DIENSTAG, J., GANEM, D. & VARMUS, H. (1989a). Detection of antibodies against hepatitis B virus polymerase antigen in hepatitis B virus-infected patients. *Hepatology* **10**, 332-335.

CHANG, L-J., PRYCIAK, P., GANEM, D. & VARMUS, H.E. (1989b). Biosynthesis of the reverse transcriptase of hepatitis B viruses involves de novo translational initiation not ribosomal frameshifting. *Nature, London* **337**, 364-368.

CHEN, H.-S., KANEKO, S., GIRONES, R., ANDERSON, R.W., HORNBUCKLE, W.E., TENNANT, B.C., COTE, P.J., GERIN, J.L., PURCELL, R.H. & MILLER, R.H.

(1993). The woodchuck hepatitis virus X gene is important for establishment of virus infection in woodchucks. *Journal of Virology* 67, 1218-1226.

CHEN, P-J., CHEN, C.-R., SUNG, J-L. & CHEN, D-S. (1989). Identification of a doubly spliced viral transcript joining the separated domains for putative protease and reverse transcriptase of hepatitis B virus. *Journal of Virology* 63, 4165-4171.

CHENG, K-C. & MOSS, B. (1987). Selective synthesis and secretion of particles composed of the hepatitis B virus middle surface protein directed by a recombinant vaccinia virus: Induction of antibodies to Pre-S and S epitopes. *Journal of Virology* 61, 1286-1290.

CHISAKA, O., ARAKI, K., OCHIYA, T., TSURIMOTO, T., HIRANYAWASITTE-ATTATIPPAHOLKUN, W., YANAIHARA, N. & MATSUBARA, K. (1987). Purification of hepatitis B virus gene X product synthesized in Escherichia coli and its detection in a human hepatoblastoma cell line producing hepatitis B virus. *Gene* 60, 183-189.

CHISARI, F.V., FILIPPI, P., BURAS, J., MCLACHLAN, A., POPPER, H., PINKERT, C.A., PALMITER, R.D. & BRINSTER, R.L. (1987). Structural and pathological effects of synthesis of hepatitis B virus large envelope polypeptide in transgenic mice. *Proceedings of the National Academy of Sciences. U.S.A.* 84, 6909-6913.

CHONG-JIN, O. & JEAK-LING, D. (1984). Hepatitis B virus in lymphocytes of seronegative carriers. *Lancet* 1, 395-396.

CHOUARD, T., BLUMENFELD, M., BACH, I., VANDEKERCKHOVE, J., CEREGHINI, S. & YANIV, M. (1990). A distal dimerization domain is essential for DNA-binding by the atypical HNF1 homeodomain. *Nucleic Acids Research* 18, 5853-5863.

CLEARY, M.A., STERN, S., TANAKA, M. & HERR, W. (1993). Differential positive control by Oct-1 and Oct-2: Activation of a transcriptionally silent motif through Oct-1 and VP16 corecruitment. *Genes and Development* 7, 72-83.

CLERC, R.G., CORCORAN, L.M., LEBOWITZ, J.H., BALTIMORE, D. & SHARP, P.A. (1988). The B-cell-specific Oct-2 protein contains POU box- and homeo box-type domains. *Genes and Development* 2, 1570-1581.

COLGROVE, R., SIMON, G. & GANEM, D. (1989). Transcriptional activation of homologous and heterologous genes by the hepatitis B virus X gene product in cells permissive for viral replication. *Journal of Virology* 63, 4019-4026.

COLUCCI, G., JACOBSON, L., LALATTA, F., VERNANCE, S., WAKSAL, S.D. & WEKSLER, M.E. (1986). Detection of hepatitis B virus DNA, RNA and surface antigen in mononuclear cells from patients with type B hepatitis. *Clinical Research* 34, 648A.

COUREY, A.J. & TJIAN, R. (1988). Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. *Cell* 55, 887-898.

COURSAGET, P., BOURDIL, C., ADAMOWICZ, P., CHOTARD, J., MAR, I.D., YVONNET, B., MEVELEC, M.N., BARRES, J.L., N'DOYE, R. & CHIRON, J.P. (1987). HBsAG positive reactivity in man not due to hepatitis B virus. *Lancet* 2, 1354-1358.

COURTOIS, G., MORGAN, J.G., CAMPBELL, L.A., FOUREL, G. & CRABTREE, G.R. (1987). Interaction of a liver-specific nuclear factor with the fibrinogen and α_1-antitrypsin promoters. *Science* 238, 688-692.

COURTOIS, G., BAUMHUETER, S. & CRABTREE, G.R. (1988). Purified hepatocyte nuclear factor 1 interacts with a family of hepatocyte-specific promoters. *Proceedings of the National Academy of Sciences, U.S.A.* 85, 7937-7941.

CURRAN, T. & FRANZA, B.R.,JR. (1988). Fos and Jun: the AP-1 connection. *Cell* 55, 395-397.

DANE, D.S., CAMERON, C.H. & BRIGGS, M. (1970). Virus-like particles in serum of patients with Australia-antigen-associated hepatitis. *Lancet* 1, 695-698.

DE FRANCESCO, R., PASTORE, A., VECCHIO, G. & CORTESE, R. (1991). Circular dichroism study of the conformational stability of the dimerization domain of transcription factor LFB1. *Biochemistry* 30, 143-147.

DE SIMONE, V., DE MAGISTRIS, L., LAZZARO, D., GERSTNER, J., MONACI, P., NICOSIA, A. & CORTESE, R. (1991). LFB3, a heterodimer-forming homeoprotein of the LFB1 family, is expressed in specialized epithelia. *EMBO Journal* 10, 1435-1443.

DE SIMONE, V. & CORTESE, R. (1992). Transcription factors and liver-specific genes. *Biochimica et Biophysica Acta Gene Structure and Expression* **1132**, 119-126.

DE WET, J.R., WOOD, K.V., DELUCA, M., HELINSKI, D.R. & SUBRAMANI, S. (1987). Firefly luciferase gene: Structure and expression in mammalian cells. *Molecular and Cellular Biology* **7**, 725-737.

DE-MEDINA, T., FAKTOR, O. & SHAUL, Y. (1988). The S promoter of hepatitis B virus is regulated by positive and negative elements. *Molecular and Cellular Biology* **8**, 2449-2455.

DEJEAN, A., LUGASSY, C., ZAFRANI, S., TIOLLAIS, P. & BRECHOT, C. (1984). Detection of hepatitis B virus DNA in pancreas, kidney and skin of two human carriers of the virus. *Journal of General Virology* **65**, 651-655.

DEJEAN, A., BOUGUELERET, L., GRZESCHIK, K-H. & TIOLLAIS, P. (1986). Hepatitis B Virus DNA integration in a sequence homologous to v-erb-A and steroid receptor genes in a hepatocellular carcinoma. *Nature, London* **322**, 70-72.

DENG, T. & KARIN, M. (1993). JunB differs from c-Jun in its DNA-binding and dimerization domains, and represses c-Jun by formation of inactive heterodimers. *Genes and Development* **7**, 479-490.

DIENSTAG, J.L., BHAN, A.K., KLINGENSTEIN, R.J. & SAVARESE, A.M. (1982). Immunopathogenesis of liver disease associated with hepatitis B. In *Viral Hepatitis, 1981 International Symposium*, pp. 221-236. Edited by G. N. Vyas, H. J. Alter & J. E. Maynard. New York: Franklin Institute Press.

DIENSTAG, J.L. (1984). Studies of cell-mediated immunity in chronic hepatitis B virus infection: The elusive goal of virus and host antigen specificity. In *Advances in Hepatitis Research*, pp. 163-167. Edited by F. V. Chisari. New York: Masson.

DIETZMAN, D.E., HARNISCH, J.P., RAY, G., ALEXANDER, E.R. & HOLMES, K.K. (1977). Hepatitis B surface antigen (HBsAg) and antibody to HBsAg: Prevalence in homosexual and heterosexual men. *Journal of the American Medical Association* **238**, 2625-2626.

DIGNAM, J.D., LEBOVITZ, R.M. & ROEDER, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Research* **11**, 1475-1489.

DIKSTEIN, R., FAKTOR, O., BEN-LEVY, R. & SHAUL, Y. (1990). Functional organization of the hepatitis B virus enhancer. *Molecular and Cellular Biology* **10**, 3683-3689.

DOCHERTY, K. & STEINER, D.F. (1982). Post-translational proteolysis in polypeptide hormone biosynthesis. *Annual Review of Physiology* **44**, 625-638.

DRANGINIS, A.M. (1990). Binding of yeast a1 and α2 as a heterodimer to the operator DNA of a haploid-specific gene. *Nature, London* **347**, 682-685.

DUBOIS, M.F., POURCEL, C., ROUSSET, S., CHANY, C. & TIOLLAIS, P. (1980). Excretion of hepatitis B surface antigen particles from mouse cells transformed with cloned viral DNA. *Proceedings of the National Academy of Sciences, U.S.A.* **77**, 4549-4553.

DUDLEY, F.J., FOX, R.A. & SHERLOCK, S. (1972). Cellular immunity and hepatitis-associated, Australia antigen liver disease. *Lancet* **1**, 723-726.

DYNLACHT, B.D., HOEY, T. & TJIAN, R. (1991). Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell* **66**, 563-576.

EBLE, B.E., LINGAPPA, V.R. & GANEM, D. (1986). Hepatitis B surface antigen: an unusual secreted protein initially synthesized as a transmembrane polypeptide. *Molecular and Cellular Biology* **6**, 1454-1463.

EBLE, B.E., MACRAE, D.R., LINGAPPA, V.R. & GANEM, D. (1987). Multiple topogenic sequences determine the transmembrane orientation of hepatitis B surface antigen. *Molecular and Cellular Biology* **7**, 3591-3601.

EBLE, B.E., LINGAPPA, V.R. & GANEM, D. (1990). The N-terminal (pre-S2) domain of a hepatitis B virus surface glycoprotein is translocated across membranes by downstream signal sequences. *Journal of Virology* **64**, 1414-1419.

ECHEVARRIA, J.M., LEON, P., DOMINGO, C.J. & LOPEZ, J.A. (1988). Atypical hepatitis B virus in Spain? *Lancet* **2**, 1315-1316.

EDMAN, J.C., GRAY, P., VALENZUELA, P., RALL, L.B. & RUTTER, W.J. (1980). Integration of hepatitis B virus sequences and their expression in a human hepatoma cell. *Nature, London* **286**, 535-538.

ELFASSI, E., ROMET-LEMONNE, J.L., ESSEX, M., FRANCES-MCLANE, M. & HASELTINE, W.A. (1984). Evidence of extrachromosomal forms of hepatitis B viral DNA in a bone marrow culture obtained from a patient recently infected with hepatitis B virus. *Proceedings of the National Academy of Sciences, U.S.A.* **81**, 3526-3528.

ELFASSI, E., HASELTINE, .W.A. & DIENSTAG, J.L. (1986). Detection of hepatitis B virus X product using an open reading frame Escherichia coli expression vector. *Proceedings of the National Academy of Sciences, U.S.A.* **83**, 2219-2222.

ELFASSI, E. (1987). Broad specificity of the hepatitis B enhancer function. *Virology* **160**, 259-262.

ENDERS, G.H., GANEM, D. & VARMUS, H.E. (1987). 5'-terminal sequences influence the segregation of ground squirrel hepatitis virus RNAs into polyribosomes and viral core particles. *Journal of Virology* **61**, 35-41.

FAKTOR, O., DE MEDINA, T. & SHAUL, Y. (1988). Regulation of hepatitis B virus S gene promoter in transfected cell lines. *Virology* **162**, 362-368.

FAKTOR, O. & SHAUL, Y. (1990). The identification of hepatitis B virus X gene responsive elements reveals functional similarity of X and HTLV-I tax. *Oncogene* **5**, 867-872.

FARZA, H., HADCHOUEL, M., SCOTTO, J., TIOLLAIS, P., BABINET, C. & POURCEL, C. (1988). Replication and gene expression of hepatitis B virus in a transgenic mouse that contains the complete viral genome. *Journal of Virology* **62**, 4144-4152.

FEAVER, W.J., GILEADI, O. & KORNBERG, R.D. (1991). Purifiction and characterization of yeast RNA polymerase II transcription factor b. *Journal of Biological Chemistry* **266**, 19000-19005.

FEITELSON, M.A., MILLMAN, I., HALBHERR, T., SIMMONS, H. & BLUMBERG, B.S. (1986). A newly identified hepatitis B type virus in tree squirrels. *Proceedings of the National Academy of Sciences, U.S.A.* **83**, 2233-2237.

FEITELSON, M.A., MILLMAN, I., DUNCAN, G.D. & BLUMBERG, B.S. (1988). Presence of antibodies to the polymerase gene product(s) of hepatitis B and woodchuck hepatitis virus in natural and experimental infections. *Journal of Medical Virology* **24**, 121-136.

FEUERMAN, M.H., GODBOUT, R., INGRAM, R.S. & TILGHMAN, S.M. (1989). Tissue-specific transcription of the mouse α-fetoprotein gene promoter is dependent on HNF-1. *Molecular and Cellular Biology* **9**, 4204-4212.

FINNEY, M., RUVKUN, G. & HORVITZ, H.R. (1988). The C. elegans cell lineage and differentiation gene unc-86 encodes a protein with a homeodomain and extended similarity to transcription factors. *Cell* **55**, 757-769.

FINNEY, M. (1990). The homeodomain of the transcription factor LF-B1 has a 21 amino acid looop between helix 2 and helix 3. *Cell* **60**, 5-6.

FRAIN, M., SWART, G., MONACI, P., NICOSIA, A., STAMPFIL, S., FRANK, R. & CORTESE, R. (1989). The liver-specific transcription factor LF-B1 contains a highly diverged homeobox DNA binding domain. *Cell* **59**, 145-157.

GALIBERT, F., MANDART, E., FITOUSSI, F., TIOLLAIS, P. & CHARNAY, P. (1979). Nucleotide sequence of the hepatitis B virus genome (subtype ayw) cloned in E. coli. *Nature, London* **281**, 646-650.

GANEM, D. (1982). Persistent infection of humans with hepatitis B virus: mechanisms and consequences. *Reviews of Infectious Diseases* **4**, 1026-1047.

GANEM, D. & VARMUS, H.E. (1987). The molecular biology of the hepatitis B viruses. *Annual Review of Biochemistry* **56**, 651-693.

GARCIA, P.D., OU, J-H., RUTTER, W.J. & WALTER, P. (1988). Targeting of the hepatitis B virus precore protein to the endoplasmic reticulum membrane: After signal peptide cleavage translocation can be aborted and the product released into the cytoplasm. *Journal of Cell Biology* **106**, 1093-1104.

GEHRING, W.J. (1987). Homeo boxes in the study of development. *Science* **236**, 1245-1252.

GERBER, M.A., SELLS, M.A., CHEN, M-L., THUNG, S.N., TABIBZADEH, S.S., HOOD, A. & ACS, G. (1988). Morphologic, immunohistochemical, and ultrastructural

studies of the production of hepatitis B virus in vitro. *Laboratory Investigation* **59**, 173-180.

GERLICH, W.H., GOLDMANN, U., MULLER, R., STIBBE, W. & WOLFF, W. (1982). Specificity and localization of the hepatitis B virus associated protein kinase. *Journal of Virology* **42**, 761-766.

GERLICH, W.H. & ROBINSON, W.S. (1980). Hepatitis B virus contains protein attached to the 5'terminus of its complete DNA strand. *Cell* **21**, 801-809.

GILEADI, O., FEAVER, W.J. & KORNBERG, R.D. (1992). Cloning of a subunit of yeast RNA polymerase II transcription factor b and CTD kinase. *Science* **257**, 1389-1392.

GILL, G. & PTASHNE, M. (1988). Negative effect of the transcriptional activator GAL4. *Nature, London* **334**, 721-724.

GINIGER, E. & PTASHNE, M. (1987). Transcription in yeast activated by a putative amphipathic α helix linked to a DNA binding unit. *Nature, London* **330**, 670-672.

GLANVILLE, N., DURNAM, D.M. & PALMITER, R.D. (1981). Structure of mouse metallothionein-I gene and its mRNA. *Nature, London* **292**, 267-269.

GORMAN, C.M., MOFFAT, L.F. & HOWARD, B.H. (1982). Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Molecular and Cellular Biology* **2**, 1044-1051.

GOUGH, N.M. (1983). Core and e antigen synthesis in rodent cells transformed with hepatitis B virus DNA is associated with greater than genome length viral messenger RNAs. *Journal of Molecular Biology* **165**, 683-699.

GOUTTE, C. & JOHNSON, A.D. (1988). a1 protein alters the DNA binding specificity of α2 repressor. *Cell* **52**, 875-882.

GOWANS, E.J., BURRELL, C.J., JILBERT, A.R. & MARMION, B.P. (1983). Patterns of single- and double-stranded hepatitis B virus DNA and viral antigen accumulation in infected liver cells. *Journal of General Virology* **64**, 1229-1239.

GOWANS, E.J., BURRELL, C.J., JILBERT, A.R. & MARMION, B.P. (1985). Cytoplasmic (but not nuclear) hepatitis B virus (HBV) core antigen reflects HBV DNA synthesis at the level of the infected hepatocyte. *Intervirology* **24**, 220-225.

GRAHAM, F.L. & VAN DER EB, A.J. (1973). A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* **52**, 456-467.

GRIPON, P., DIOT, C., THEZE, N., FOUREL, I., LOREAL, O., BRECHOT, C. & GUGUEN-GUILLOUZO, C. (1988). Hepatitis B virus infection of adult human hepatocytes cultured in the presence of dimethyl sulfoxide. *Journal of Virology* **62**, 4136-4143.

GUO, W., BELL, K.D. & OU, J.-H. (1991). Characterization of the hepatitis B virus EnhI enhancer and X promoter complex. *Journal of Virology* **65**, 6686-6692.

GUST, I.D., CROSS, G., KALDOR, J. & FERRIS, A.A. (1970). Virus-like particles in Australia-antigen-associated hepatitis. *Lancet* **1**, 953.

GUST, I.D., BURRELL, C.J., COULEPIS, A.G., ROBINSON, W.S. & ZUCKERMAN, A.J. (1986). Taxonomic classification of human hepatitis B virus. *Intervirology* **25**, 14-29.

HADCHOUEL, M., PASQUINELLI, C., FOURNIER, J.G., HUGON, R.N., SCOTTO, J., BERNARD, O. & BRECHOT, C. (1988). Detection of mononuclear cells expressing hepatitis B virus in peripheral blood from HBsAg positive and negative patients by in situ hybridisation. *Journal of Medical Virology* **24**, 27-32.

HAM, J., STEGER, G. & YANIV, M. (1992). How do eukaryotic activator proteins stimulate the rate of transcription by RNA polymerase II. *FEBS Letters* **307**, 81-86.

HARDON, E.M., FRAIN, M., PAONESSA, G. & CORTESE, R. (1988). Two distinct factors interact with the promoter regions of several liver-specific genes. *EMBO Journal* **7**, 1711-1719.

HEERMANN, K.H., GOLDMANN, U., SCHWARTZ, W., SEYFFARTH, T., BAUMGARTEN, H. & GERLICH, W.H. (1984). Large surface proteins of hepatitis B virus containing the Pre-s sequence. *Journal of Virology* **52**, 396-402.

HENRY, N.L., SAYRE, M.H. & KORNBERG, R.D. (1992). Purification and characterization of yeast RNA polymerase II general initiation factor g. *Journal of Biological Chemistry* **267**, 23388-23392.

HERR, W., STURM, R.A., CLERC, R.G., CORCORAN, L.M., BALTIMORE, D., SHARP, P.A., INGRAHAM, H.A., ROSENFELD, M.G., FINNEY, M., RUVKUN, G.

& HORVITZ, H.R. (1988). The POU domain: a large conserved region in the mammaliam pit-1, oct-1, oct-2, and Caenorhabditis elegans unc-86 gene products. *Genes and Development* 2, 1513-1516.

HIRSCH, R.C., LAVINE, J.E., CHANG, L-J., VARMUS, H.E. & GANEM, D. (1990). Polymerase gene products of hepatitis B viruses are required for genomic RNA packaging as well as for reverse transcription. *Nature, London* 344, 552-555.

HIRSCHMAN, R.J., SHULMAN, R., BARKER, L.F. & SMITH, K.O. (1969). Virus-like particles in sera of patients with infectious and serum hepatitis. *Journal of the American Medical Association* 208, 1667-1670.

HIRSCHMAN, S.Z. (1984). Replication of hepatitis B virus in cell culture systems. In *Advances in Hepatitis Research*, pp. 54-61. Edited by F. V. Chisari. New York: Masson.

HOEY, T., DYNLACHT, B.D., PETERSON, M.G., PUGH, B.F. & TJIAN, R. (1990). Isolation and characterization of the Drosophila gene encoding the TATA box binding protein, TFIID. *Cell* 61, 1179-1186.

HONIGWACHS, J., FAKTOR, O., DIKSTEIN, R., SHAUL, Y. & LAUB, O. (1989). Liver-specific expression of hepatitis B virus is determined by the combined action of the core gene promoter and the enhancer. *Journal of Virology* 63, 919-924.

HOOFNAGLE, J.H., DUSHEIKO, G.M., SEEFF, L.B., JONES, E.A., WAGGONER, J.G. & BALES, Z.B. (1981). Seroconversion from hepatitis B e antigen to antibody in chronic type B hepatitis. *Annals of Internal Medicine* 94, 744-748.

HOPE, I.A. & STRUHL, K. (1986). Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell* 46, 885-894.

HRUSKA, J.F., CLAYTON, D.A., RUBENSTEIN, J.L.R. & ROBINSON, W.S. (1977). Structure of hepatitis B Dane particle DNA before and after the Dane particle DNA polymerase reaction. *Journal of Virology* 21, 666-672.

HRUSKA, J.F. & ROBINSON, W.S. (1977). The proteins of hepatitis B Dane particle cores. *Journal of Medical Virology* 1, 119-131.

HSIUNG, N., FITTS, R., WILSON, S., MILNE, A. & HAMER, D.H. (1984). Efficient production of hepatitis B surface antigen using a bovine papilloma virus-metallothionein vector. *Journal of Molecular and Applied Genetics* 2, 497-506.

HUANG, S-N. & GROH, V. (1973). Immunoagglutination electron microscopic study on virus-like particles and Australia antigen in liver tissue. *Laboratory Investigation* **29**, 353-366.

IMAZEKI, F., YAGINUMA, K., OMATA, M., OKUDA, K., KOBAYASHI, M. & KOIKE, K. (1987). RNA transcripts of hepatitis B virus in hepatocellular carcinoma. *Hepatology* **7**, 753-757.

INGRAHAM, H.A., CHEN, R., MANGALAM, H.J., ELSHOLTZ, H.P., FLYNN, S.E., LIN, C.R., SIMMONS, D.M., SWANSON, L. & ROSENFELD, M.G. (1988). A tissue-specific transcription factor containing a homeodomain specifies a pituitary phenotype. *Cell* **55**, 519-529.

INGRAHAM, H.A., FLYNN, S.E., VOSS, J.W., ALBERT, V.R., KAPILOFF, M.S., WILSON, L. & ROSENFELD, M.G. (1990). The POU-specific domain of Pit-1 is essential for sequence-specific, high affinity DNA binding and DNA-dependent Pit-1-Pit-1 interactions. *Cell* **61**, 1021-1033.

JAMEEL, S. & SIDDIQUI, A. (1986). The human hepatitis B virus enhancer requires trans-acting cellular factor(s) for activity. *Molecular and Cellular Biology* **6**, 710-715.

JEAN-JEAN, O., LEVRERO, M., WILL, H., PERRICAUDET, M. & ROSSIGNOL, J-M. (1989). Expression mechanism of the hepatitis B virus (HBV) C gene and biosynthesis of HBe antigen. *Virology* **170**, 99-106.

JOHNSON, A.D. & HERSKOWITZ, I. (1985). A repressor (MATα2 product) and its operator control expression of a set of cel type specific genes in yeast. *Cell* **42**, 237-247.

JOKELAINEN, P.T., KROHN, K., PRINCE, A.M. & FINLAYSON, N.D.C. (1970). Electron microscopic observations on virus-like particles associated with SH antigen. *Journal of Virology* **6**, 685-689.

JUNKER, M., GALLE, P. & SCHALLER, H. (1987). Expression and replication of the hepatitis B virus genome under foreign promoter control. *Nucleic Acids Research* **15**, 10117-10132.

KADONAGA, J.T., COUREY, A.J., LADIKA, J. & TJIAN, R. (1988). Distinct regions of Sp1 modulate DNA binding and transcriptional activiation. *Science* **242**, 1566-1570.

KAMIMURA, T., YOSHIKAWA, A., ICHIDA, F. & SASAKI, H. (1981). Electron microscopic studies of Dane particles in hepatocytes with special reference to intracellular development of Dane particles and their relation with HBeAg in serum. *Hepatology* **1**, 392-397.

KANEKO, S. & MILLER, R.H. (1988). X-region-specific transcript in mammalian hepatitis B virus-infected liver. *Journal of Virology* **62**, 3979-3984.

KAO, C.C., LIEBERMAN, P.M., SCHMIDT, M.C., ZHOU, Q., PEI, R. & BERK, A.J. (1990). Cloning of a transcriptionally active human TATA binding factor. *Science* **248**, 1646-1650.

KAPLAN, P.M., GREENMAN, R.L., GERIN, J.L., PURCELL, R.H. & ROBINSON, W.S. (1973). DNA polymerase associated with human hepatitis B antigen. *Journal of Virology* **12**, 995-1005.

KARPEN, S., BANERJEE, R., ZELENT, A., PRICE, P. & ACS, G. (1988). Identification of protein-binding sites in the hepatitis B virus enhancer and core promoter domains. *Molecular and Cellular Biology* **8**, 5159-5165.

KATAYAMA, K., HAYASHI, N., SASAKI, Y., KASAHARA, A., UEDA, K., FUSAMOTO, H., SATO, N., CHISAKA, O., MATSUBARA, K. & KAMADA, T. (1989). Detection of hepatitis B virus X gene protein and antibody in type B chronic liver disease. *Gastroenterology* **97**, 990-998.

KAY, A., MANDART, E., TREPO, C. & GALIBERT, F. (1985). The HBV HBx gene expressed in E. coli is recognised by sera from hepatitis patients. *EMBO Journal* **4**, 1287-1292.

KELLEHER, R.J.,III, FLANAGAN, P.M. & KORNBERG, R.D. (1990). A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. *Cell* **61**, 1209-1215.

KILLEEN, M., COULOMBE, B. & GREENBLATT, J. (1992). Recombinant TBP, transcription factor IIB, and RAP30 are sufficient for promoter recognition by mammalian RNA polymerase II. *Journal of Biological Chemistry* **267**, 9463-9466.

KISSINGER, C.R., LIU, B., MARTIN-BLANCO, E., KORNBERG, T.B. & PABO, C.O. (1990). Crystal structure of an engrailed homeodomain-DNA complex at 2.8Å resolution: a framework for understanding homeodomain-DNA interactions. *Cell* **63**, 579-590.

KLINKERT, M-Q., THEILMANN, L., PFAFF, E. & SCHALLER, H. (1986). Pre-s1 antigens and antibodies early in the course of acute hepatitis B virus infection. *Journal of Virology* **58**, 522-525.

KNOWLES, B.B., HOWE, C.C. & ADEN, D.P. (1980). Human hpatocellular carcinoma cell lines secrete the major plasma proteins and hepatitis B surface antigen. *Science* **209**, 497-499.

KOBAYASHI, M. & KOIKE, K. (1984). Complete nucleotide sequence of hepatitis B virus DNA of subtype adr and its conserved gene organization. *Gene* **30**, 227-232.

KOMAI, K., KAPLAN, M. & PEEPLES, M.E. (1988). The vero cell receptor for the hepatitis B virus small S protein is a sialoglycoprotein. *Virology* **163**, 629-634.

KORBA, B.E., WELLS, F., TENNANT, B.C., YOAKUM, G.H., PURCELL, R.H. & GERIN, J.L. (1986). Hepadnavirus infection of peripheral blood lymphocytes in vivo: Woodchuck and chimpanzee models of viral hepatitis. *Journal of Virology* **58**, 1-8.

KRISTIE, T.M. & SHARP, P.A. (1990). Interactions of the Oct-1 POU subdomains with specific DNA sequences and with the HSV α-*trans*-activator protein. *Genes and Development* **4**, 2383-2396.

KUO, C.J., CONLEY, P.B., HSIEH, C.-L., FRANCKE, U. & CRABTREE, G.R. (1990). Molecular cloning, functional expression, and chromosomal localization of mouse hepatocyte nuclear factor 1. *Proceedings of the National Academy of Sciences, U.S.A.* **87**, 9838-9842.

KUO, C.J., CONLEY, P.B., CHEN, L., SLADEK, F.M., DARNELL, J.E.,JR. & CRABTREE, G.R. (1992). A transcriptional hierarchy involved in mammalian cell-type specification. *Nature, London* **355**, 457-461.

LAMELIN, J-P. & TREPO, C. (1990). The hepatitis B virus and the peripheral blood mononuclear cells: A brief review. *Journal of Hepatology* **10**, 120-124.

LANDERS, T.A., GREENBERG, H.B. & ROBINSON, W.S. (1977). Structure of hepatitis B Dane particle DNA and nature of the endogenous DNA polymerase reaction. *Journal of Virology* **23**, 368-376.

LANDT, O., GRUNERT, H.P. & HAHN, U. (1990). A general method for rapid site-directed mutagenesis using the polymerase chain reaction. *Gene* **96**, 125-128.

LANGLEY, K.E., EGAN, K.M., BARENDT, J.M., PARKER, C.G. & BITTER, G.A. (1988). Characterization of purified hepatitis B surface antigen containing pre-S(2) epitopes expressed in Saccharomyces cerevisiae. *Gene* **67**, 229-245.

LAUER, J.L., VANDRUNEN, N.A., WASHBURN, J.W. & BALFOUR, H.H. (1979). Transmission of hepatitis B virus in clinical laboratory areas. *Journal of Infectious Diseases* **140**, 513-516.

LAUGHON, A. & SCOTT, M.P. (1984). Sequence of a Drosophila segmentation gene: protein structure homology with DNA-binding proteins. *Nature, London* **310**, 25-31.

LEE, A.K.Y., IP, H.M.H. & WONG, V.C.W. (1978). Mechanisms of maternal-fetal transmission of hepatitis B virus. *Journal of Infectious Diseases* **138**, 668-671.

LEVRERO, M., JEAN-JEAN, O., BALSANO, C., WILL, H. & PERRICAUDET, M. (1990). Hepatitis B virus (HBV) X gene expression in human cells and anti-HBx antibodies detection in chronic HBV infection. *Virology* **174**, 299-304.

LEWIS, T.L., ALTER, H.J., CHALMERS, T.C., HOLLAND, P.V., PURCELL, R.H., ALLING, D.W., YOUNG, D., FRENKEL, L.D., LEE, S.L. & LAMSON, M.E. (1973). A comparison of the frequency of hepatitis-B antigen and antibody in hospital and nonhospital personnel. *New England Journal of Medicine* **289**, 647-651.

LIANG, X.-H., STEMLER, M., WILL, H., BRAUN, R., TANG, Z.-Y. & SCHRODER, C.H. (1988). Low incidence and high titers of antibodies to hepatitis B virus X-protein in sera of chinese patients with hepatocellular carcinoma. *Journal of Medical Virology* **25**, 329-337.

LICHTER, E.A. (1969). Chimpanzee antibodies to Australia antigen. *Nature, London* **224**, 810-811.

LICHTSTEINER, S. & SCHIBLER, U. (1989). A glycosylated liver-specific transcription factor stimulates transcription of the albumin gene. *Cell* **57**, 1179-1187.

LIEN, J-M., ALDRICH, C.E. & MASON, W.S. (1986). Evidence that a capped oligoribonucleotide is the primer for duck hepatitis B virus plus-strand DNA synthesis. *Journal of Virology* 57, 229-236.

LIN, Y.-S., HA, I., MALDONADO, E., REINBERG, D. & GREEN, M.R. (1991). Binding of general transcription factor TFIIB to an acidic activating region. *Nature, London* 353, 569-571.

LINNEY, E. & DONERLY, S. (1983). DNA fragments from F9 PyEC mutants increase expression of heterologous genes in transfected F9 cells. *Cell* 35, 693-699.

LIU, F. & GREEN, M.R. (1990). A specific member of the ATF transcription factor family can mediate transcription activation by the adenovirus E1a protein. *Cell* 61, 1217-1224.

LONDON, W.T., MILLMAN, I., SUTTNICK, A.I. & BLUMBERG, B.S. (1970). Transmission, replication, and passage of Australia antigen in African green monkeys. *Clinical Research* 18, 536.

LONDON, W.T., ALTER, H.J., LANDER, J. & PURCELL, R.H. (1972). Serial transmission in rhesus monkeys of an agent related to hepatitis-associated antigen. *Journal of Infectious Diseases* 125, 382-389.

LOPEZ-CABRERA, M., LETOVSKY, J., HU, K.-Q. & SIDDIQUI, A. (1990). Multiple liver-specific factors bind to the hepatitis B virus core/pregenomic promoter: Trans-activation and repression by CCAAT/enhancer binding protein. *Proceedings of the National Academy of Sciences, U.S.A.* 87, 5069-5073.

LOPEZ-CABRERA, M., LETOVSKY, J., HU, K.-Q. & SIDDIQUI, A. (1991). Transcriptional factor C/EBP binds to and transactivates the enhancer element II of the hepatitis B virus. *Virology* 183, 825-829.

LUCITO, R. & SCHNEIDER, R.J. (1992). Hepatitis B virus X protein activates transcription factor NF-kappaB without a requirement for protein kinase C. *Journal of Virology* 66, 983-991.

LUTWICK, L.I. & ROBINSON, W.S. (1977). DNA synthesized in the hepatitis B Dane particle DNA polymerase reaction. *Journal of Virology* 21, 96-104.

MA, J. & PTASHNE, M. (1987). Deletion analysis of GAL4 defines two transcriptional activating segments. *Cell* 48, 847-853.

MACGREGOR, P.F., ABATE, C. & CURRAN, T. (1990). Direct cloning of leucine zipper proteins: Jun binds cooperatively to the CRE with CRE-BP1. *Oncogene* 5, 451-458.

MACHIDA, A., KISHIMOTO, S., OHNUMA, H., BABA, K., ITO, Y., MIYAMOTO, H., FUNATSU, G., ODA, K., USUDA, S., TOGAMI, S., NAKAMURA, T., MIYAKAWA, Y. & MAYUMI, M. (1984). A polypeptide containing 55 amino acid residues coded by the Pre-S region of hepatitis B virus deoxyribonucleic acid bears the receptor for polymerized human as well as chimpanzee albumins. *Gastroenterology* 86, 910-918.

MACK, D.H., BLOCH, W., NATH, N. & SNINSKY, J.J. (1988). Hepatitis B virus particles contain a polypeptide encoded by the largest open reading frame: A putative reverse transcriptase. *Journal of Virology* 62, 4786-4790.

MACNAB, G.M., ALEXANDER, J.J., LECATSAS, G., BEY, E.M. & URBANOWICZ, J.M. (1976). Hepatitis B surface antigen produced by a human hepatoma cell line. *British Journal of Cancer* 34, 509-515.

MAGUIRE, H.F., HOEFFLER, J.P. & SIDDIQUI, A. (1991). HBV X protein alters the DNA binding specificity of CREB and ATF-2 by protein-protein interactions. *Science* 252, 842-844.

MAIRE, P., WUARIN, J. & SCHIBLER, U. (1989). The role of cis-acting promoter elements in tissue-specific albumin gene expression. *Science* 244, 343-346.

MALDONADO, E., HA, I., CORTES, P., WEIS, L. & REINBERG, D. (1990). Factors involved in specific transcription by mammalian RNA polymerase II: role of trnascription factors IIA, IID, and IIB during the formation of a transcription-competent complex. *Molecular and Cellular Biology* 10, 6335-6347.

MARION, P.L., OSHIRO, L.S., REGNERY, D.C., SCULLARD, G.H. & ROBINSON, W.S. (1980). A virus in Beechey ground squirrels that is related to hepatitis B virus of humans. *Proceedings of the National Academy of Sciences, U.S.A.* 77, 2941-2945.

MARKENSON, J.A., GERETY, R.J., HOOFNAGLE, J.H. & BARKER, L.F. (1975). Effects of cyclophosphamide on hepatitis B virus infection and challenge in chimpanzees. *Journal of Infectious Diseases* **131**, 79-87.

MASON, W.S., SEAL, G. & SUMMERS, J. (1980). Virus of Pekin ducks with structural and biological relatedness to human hepatitis B virus. *Journal of Virology* **36**, 829-836.

MAYNARD, J.E., HARTWELL, W.V. & BERQUIST, K.R. (1971). Hepatitis-associated antigen in chimpanzees. *Journal of Infectious Diseases* **123**, 660-664.

MAYNARD, J.E., BERQUIST, K.R., KRUSHAK, D.H. & PURCELL, R.H. (1972). Experimental infection of chimpanzees with the virus of hepatitis B. *Nature, London* **237**, 514-515.

MCGINNIS, W., GARBER, R.L., WIRZ, J., KUROIWA, A. & GEHRING, W.J. (1984a). A homologous protein-coding sequence in Drosophila homeotic genes and its conservation in other metazoans. *Cell* **37**, 403-408.

MCGINNIS, W., LEVINE, M.S., HAFEN, E., KUROIWA, A. & GEHRING, W.J. (1984b). A conserved DNA sequence in homoeotic genes of the Drosophila Antennapedia and bithorax complexes. *Nature, London* **308**, 428-433.

MCKAY, D.B. & STEITZ, T.A. (1981). Structure of catabolite gene activator protein at 2.9Å resolution suggests binding to left-handed B-DNA. *Nature, London* **290**, 744-749.

MCLACHLAN, A., MILICH, D.R., RANEY, A.K., RIGGS, M.G., HUGHES, J.L., SORGE, J. & CHISARI, F.V. (1987). Expression of hepatitis B virus surface and core antigens: Influences of pre-S and precore sequences. *Journal of Virology* **61**, 683-692.

MENDEL, D.B., HANSEN, L.P., GRAVES, M.K., CONLEY, P.B. & CRABTREE, G.R. (1991a). HNF-1α and HNF-1β (vHNF-1) share dimerization and homeo domains, but not activation domains, and form heterodimers in vitro. *Genes and Development* **5**, 1042-1056.

MENDEL, D.B., KHAVARI, P.A., CONLEY, P.B., GRAVES, M.K., HANSEN, L.P., ADMON, A. & CRABTREE, G.R. (1991b). Characterization of a cofactor that regulates dimerization of a mammalian homeodomain protein. *Science* **254**, 1762-1767.

MENDEL, D.B. & CRABTREE, G.R. (1991). HNF-1, a member of a novel class of dimerizing homeodomain proteins. *Journal of Biological Chemistry* **266**, 677-680.

MERMOD, N., O'NEILL, E.A., KELLY, T.J. & TJIAN, R. (1989). The proline-rich transcriptional activator of CTF/NF1 is distinct from the replication and DNA binding domain. *Cell* **58**, 741-753.

MEYER, M.E., GRONEMEYER, H., TURCOTTE, B., BOCQUEL, M.T., TASSET, D. & CHAMBON, P. (1989). Steroid hormone receptors compete for factors that mediate their enhancer function. *Cell* **57**, 433-442.

MEYERS, M.L., TREPO, L.V., NATH, N. & SNINSKY, J.J. (1986). Hepatitis B virus polypeptide X: Expression in Escherichia coli and identification of specific antibodies in sera from hepatitis B virus-infected humans. *Journal of Virology* **57**, 101-109.

MICHEL, M-L., PONTISSO, P., SOBCZAK, E., MALPIECE, Y., STREECK, R.E. & TIOLLAIS, P. (1984). Synthesis in animal cells of hepatitis B surface antigen particles carrying a receptor for polymerized human serum albumin. *Proceedings of the National Academy of Sciences, U.S.A.* **81**, 7708-7712.

MILICH, D.R., JONES, J.E., HUGHES, J.L., PRICE, J., RANEY, A.K. & MCLACHLAN, A. (1990). Is a function of the secreted hepatitis B e antigen to induce immunologic tolerance: a transgenic model. *Proceedings of the National Academy of Sciences, U.S.A.* **in press**,

MILLER, R.H., TRAN, C-T. & ROBINSON, W.S. (1984). Hepatitis B virus particles of plasma and liver contain viral DNA-RNA hybrid molecules. *Virology* **139**, 53-63.

MILLER, R.H. & ROBINSON, W.S. (1984). Hepatitis B virus DNA forms in nuclear and cytoplasmic fractions of infected human liver. *Virology* **137**, 390-399.

MITCHELL, P.J. & TJIAN, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245**, 371-378.

MOLNAR-KIMBER, K.L., JAROCKI-WITEK, V., DHEER, S.K., VERNON, S.K., CONLEY, A.J., DAVIS, A.R. & HUNG, P.P. (1988). Distinctive properties of the hepatitis B virus envelope proteins. *Journal of Virology* **62**, 407-416.

MONACI, P., NICOSIA, A. & CORTESE, R. (1988). Two different liver-specific factors stimulate in vitro transcription from the human α1-antitrypsin promoter. *EMBO Journal* **7**, 2075-2087.

MORIARTY, A.M, HOYER, B.H., SHIH, J.W-K., GERIN, J.L. & HAMER, D.H. (1981). Expression of the hepatitis B virus surface antigen gene in cell culture by using a simian virus 40 vector. *Proceedings of the National Academy of Sciences, U.S.A.* 78, 2606-2610.

MORIARTY, A.M., ALEXANDER, H., LERNER, R.A. & THORNTON, G.B. (1985). Antibodies to peptides detect new hepatitis B antigen: serological correlation with hepatocellular carcinoma. *Science* 227, 429-433.

MULLER, M., AFFOLTER, M., LEUPIN, W., OTTING, G., WUTHRICH, K. & GEHRING, W.J. (1988). Isolation and sequence-specific DNA binding of the Antennapedia homeodomain. *EMBO Journal* 7, 4299-4304.

NAKABAYASHI, H., TAKETA, K., MIYANO, K., YAMANE, T. & SATO, J. (1982). Growth of human hepatoma cell lines with differentiated functions in chemically defined medium. *Cancer Research* 42, 3858-3863.

NAKAO, K., MIYAO, Y., OHE, Y. & TAMAOKI, T. (1989). Involvement of an AFP1-binding site in cell-specific transcription of the pres-S1 region of the human hepatitis B virus surface antigen gene. *Nucleic Acids Research* 17, 9833-9842.

NEURATH, A.R., KENT, S.B.H., STRICK, N., TAYLOR, P. & STEVENS, C.E. (1985). Hepatitis B virus contains pre-S gene-encoded domains. *Nature, London* 315, 154-156.

NEURATH, A.R., KENT, S.B.H. & STRICK, N. (1986a). Detection of antiviral antibodies with predetermined specificity using synthetic peptide-β lactamase conjugates: Application to antibodies specific for the preS region of the hepatitis B virus envelope proteins. *Journal of General Virology* 67, 453-461.

NEURATH, A.R., KENT, S.B.H., STRICK, N. & PARKER, K. (1986b). Identification and chemical synthesis of a host cell receptor binding site on hepatitis B virus. *Cell* 46, 429-436.

NICOSIA, A., MONACI, P., TOMEI, L., DE FRANCESCO, R., NUZZO, M., STUNNENBERG, H. & CORTESE, R. (1990). A myosin-like dimerization helix and an extra-large homeodomain are essential elements of the tripartite DNA binding structure of LFB1. *Cell* 61, 1225-1236.

NISHIOKA, K., LEVIN, A.G. & SIMONS, M.J. (1975). Hepatitis B antigen, antigen subtypes, and hepatitis B antibody in normal subjects and patients with liver disease. *Bulletin of the World Health Organization* **52**, 293-300.

OCKNER, R.K., WEISIGER, R.A. & GOLLAN, J.L. (1983). Hepatic uptake of albumin-bound substances: albumin receptor concept. *American Journal of Physiology* **245**, G13-G18.

OKAMOTO, H., IMAI, M., SHIMOZAKI, M., HOSHI, Y., IIZUKA, H., GOTANDA, T., TSUDA, F., MIYAKAWA, Y. & MAYUMI, M. (1986). Nucleotide sequence of a cloned hepatitis B virus genome, subtype ayr: Comparison with genomes of the other three subtypes. *Journal of General Virology* **67**, 2305-2314.

OKAMOTO, H., WANG, Y., TANAKA, T., MACHIDA, A., MIYAKAWA, Y. & MAYUMI, M. (1993). Trans-complementation among naturally occurring deletion mutants of hepatitis B virus and integrated viral DNA for the production of viral particles with mutant genomes in hepatoma cell lines. *Journal of General Virology* **74**, 407-414.

ONO, Y., ONDA, H., SASADA, R., IGARASHI, K., SUGINO, Y. & NISHIOKA, K. (1983). The complete nucleotide sequences of the cloned hepatitis B virus DNA: subtype adr and adw. *Nucleic Acids Research* **11**, 1747-1757.

ONODERA, S., OHORI, H., YAMAKI, M. & ISHIDA, N. (1982). Electron microscopy of human hepatitis B virus cores by negative staining-carbon film technique. *Journal of Medical Virology* **10**, 147-155.

OTT, M.-O., REY-CAMPOS, J., CEREGHINI, S. & YANIV, M. (1991). vHNF1 is expressed in epithelial cells of distinct embryonic origin during development and precedes HNF1 expression. *Mechanisms of Development* **36**, 47-58.

OTTING, G., QIAN, Y.Q., BILLETER, M., MULLER, M., AFFOLTER, M., GEHRING, W.J. & WUTHRICH, K. (1990). Protein-DNA contacts in the structure of a homeodomain-DNA complex determined by nuclear magnetic resonance spectroscopy in solution. *EMBO Journal* **9**, 3085-3092.

OU, J-H. & RUTTER, W.J. (1987). Regulation of secretion of the hepatitis B virus major surface antigen by the PreS-1 protein. *Journal of Virology* **61**, 782-786.

OU, J.H. & RUTTER, W.J. (1985). Hybrid hepatitis B virus-host transcripts in a human hepatoma cell. *Proceedings of the National Academy of Sciences, U.S.A.* **82**, 83-87.

PARKER, B.A. & STARK, G.R. (1979). Regulation of simian virus 40 transcription: sensitive analysis of the RNA species present early in infections by virus or viral DNA. *Journal of Virology* **31**, 360-369.

PARVAZ, P., LAMELIN, J-P., VITVITSKI, L., BOUFFARD, P., PICHOUD, C., COVA, L. & TREPO, C. (1987). Prevalence and significance of hepatitis B virus antigens: Expression in peripheral blood mononuclear cells in chronic active hepatitis. *Clinical Immunology and Immunopathology* **43**, 1-8.

PASEK, M., GOTO, T., GILBERT, W., ZINK, B., SCHALLER, H., MACKAY, P., LEADBETTER, G. & MURRAY, K. (1979). Hepatitis B virus genes and their expression in E. coli. *Nature, London* **282**, 575-579.

PASTORE, A., DEFRANCESCO, R., BARBATO, G., MORELLI, M.A.C., MOTTA, A. & CORTESE, R. (1991). ^{1}H resonance assignment and secondary structure determination of the dimerization domain of the trnascription factor LFB1. *Biochemistry* **30**, 148-153.

PATEL, N.U., JAMEEL, S., ISOM, H. & SIDDIQUI, A. (1989). Interactions between nuclear factors and the hepatitis B virus enhancer. *Journal of Virology* **63**, 5293-5301.

PATZER, E.J., NAKAMURA, G.R. & YAFFE, A. (1984). Intracellular transport and secretion of hepatitis B virus surface antigen in mammalian cells. *Journal of Virology* **51**, 346-353.

PATZER, E.J., NAKAMURA, G.R., SIMONSEN, C.C., LEVINSON, A.D. & BRANDS, R. (1986). Intracellular assembly and packaging of hepatitis B surface antigen particles occur in the endoplasmic reticulum. *Journal of Virology* **58**, 884-892.

PEEPLES, M.E., KOMAI, K., RADEK, R. & BANKOWSKI, M.J. (1987). A cultured cell receptor for the small S protein of hepatitis B virus. *Virology* **160**, 135-142.

PERSING, D.H., VARMUS, H.E. & GANEM, D. (1985). A frameshift mutation in the pre-S region of the human hepatitis B virus genome allows production of surface antigen particles but eliminates binding to polymerized albumin. *Proceedings of the National Academy of Sciences, U.S.A.* **82**, 3440-3444.

PETERSON, D.L., ROBERTS, I.M. & VYAS, G.N. (1977). Partial amino acid sequence of two major component polypeptides of hepatitis B surface antigen. *Proceedings of the National Academy of Sciences, U.S.A.* 74, 1530-1534.

PETERSON, D.L. (1981). Isolation and characterization of the major protein and glycoprotein of hepatitis B surface antigen. *Journal of Biological Chemistry* 256, 6975-6983.

PETERSON, M.G., TANESE, N., PUGH, B.F. & TJIAN, R. (1990). Functional domains and upstream activation properties of cloned human TATA binding protein. *Science* 248, 1625-1630.

PETIT, M-A. & PILLOT, J. (1985). HBc and HBe antigenicity and DNA-binding activity of major core protein P22 in hepatitis B virus core particles isolated from the cytoplasm of human liver cells. *Journal of Virology* 53, 543-551.

PONTISSO, P., PETIT, M-A., BANKOWSKI, M.J. & PEEPLES, M.E. (1989a). Human liver plasma membranes contain receptors for the hepatitis B virus Pre-S1 region and, via polymerized human serum albumin, for the Pre-S2 region. *Journal of Virology* 63, 1981-1988.

PONTISSO, P., RUVOLETTO, M.G., GERLICH, W.H., HEERMANN, K-H., BARDINI, R. & ALBERTI, A. (1989b). Identification of an attachment site for human liver plasma membranes on hepatitis B virus particles. *Virology* 173, 522-530.

POURCEL, C., LOUIS, A., GERVAIS, M., CHENCINER, N., DUBOIS, M-F. & TIOLLAIS, P. (1982). Transcription of the hepatitis B surface antigen gene in mouse cells transformed with cloned viral DNA. *Journal of Virology* 42, 100-105.

PRICE, P.M., MOHAMAD, A., ZELENT, A., NEURATH, A.R. & ACS, G. (1988). Translational selection in the expression of the hepatitis B virus envelope proteins. *DNA* 7, 417-422.

PRINCE, A.M. (1968a). An antigen detected in the blood during the incubation period of serum hepatitis. *Proceedings of the National Academy of Sciences, U.S.A.* 60, 814-821.

PRINCE, A.M. (1968b). Relation of Australia and SH antigens. *Lancet* 2, 462-463.

PRINCE, A.M., HARGROVE, R.L., SZMUNESS, W., CHERUBIN, C.E., FONTANA, V.J. & JEFFRIES, G.H. (1970). Immunologic distinction between infectious and serum hepatitis. *New England Journal of Medicine* **282**, 987-991.

PTASHNE, M. & GANN, A.A.F. (1990). Activators and targets. *Nature, London* **346**, 329-331.

PUCK, T.T., MARCUS, P.I. & CIECIURA, S.J. (1956). Clonal growth of mammalian cells in vitro: growth characteristics of colonies from single HeLa cells with and without a "feeder" layer. *Journal of Experimental Medicine* **103**, 273-284.

PUGH, B.F. & TJIAN, R. (1990). Mechanism of transcriptional activation by Sp1: Evidence for coactivators. *Cell* **61**, 1187-1197.

PUGH, B.F. & TJIAN, R. (1991). Transcription from a TATA-less promoter requires a multisubunit TFIID complex. *Genes and Development* **5**, 1935-1945.

RADZIWILL, G., TUCKER, W. & SCHALLER, H. (1990). Mutational analysis of the hepatitis B virus P gene product: Domain structure and RNase activity. *Journal of Virology* **64**, 613-620.

RALL, L.B., STANDRING, D.N., LAUB, O. & RUTTER, W.J. (1983). Transcription of hepatitis B virus by RNA polymerase II. *Molecular and Cellular Biology* **3**, 1766-1773.

RANEY, A.K., MILICH, D.R. & MCLACHLAN, A. (1989). Characterization of hepatitis B virus major surface antigen gene transcriptional regulatory elements in differentiated hepatoma cell lines. *Journal of Virology* **63**, 3919-3925.

RANEY, A.K., MILICH, D.R., EASTON, A.J. & MCLACHLAN, A. (1990). Differentiation specific transcriptional regulation of the hepatitis B virus large surface antigen gene in human hepatoma cell lines. *Journal of Virology* **64**, 2360-2368.

RANEY, A.K., EASTON, A.J., MILICH, D.R. & MCLACHLAN, A. (1991a). Promoter-specific transactivation of hepatitis B virus transcription by a glutamine- and proline-rich domain of hepatocyte nuclear factor 1. *Journal of Virology* **65**, 5774-5781.

RANEY, A.K., MILICH, D.R. & MCLACHLAN, A. (1991b). Complex regulation of transcription from the hepatitis B virus major surface antigen promoter in human hepatoma cell lines. *Journal of Virology* **65**, 4805-4811.

RANEY, A.K., LE, H.B. & MCLACHLAN, A. (1992). Regulation of transcription from the hepatitis B virus major surface antigen promoter by the Sp1 transcription factor. *Journal of Virology* **66**, 6912-6921.

RANEY, A.K. & MCLACHLAN, A. (1991). The Biology of Hepatitis B Virus. In *Molecular biology of the hepatitis B virus*, pp. 1-37. Edited by A. McLachlan. Boca Raton, Florida: CRC Press.

RANSONE, L.J., VISVADER, J., SASSONE CORSI, P. & VERMA, I.M. (1989). Fos-Jun interaction: mutational analysis of the leucine zipper domain of both proteins. *Genes and Development* **3**, 770-781.

REALDI, G., ALBERTI, A., RUGGE, M., BORTOLOTTI, F., RIGOLI, A.M., TREMOLADA, F. & RUOL, A. (1980). Seroconversion from hepatitis B e antigen to anti-HBe in chronic hepatitis B virus infection. *Gastroenterology* **79**, 195-199.

REDEKER, A.G. & GOVINDARAJAN, S. (1988). Possible case of HBV-2 infection. *Lancet* **1**, 656.

REY-CAMPOS, J., CHOUARD, T., YANIV, M. & CEREGHINI, S. (1991). vHNF1 is a homeoprotein that activates transcription and forms heterodimers with HNF1. *EMBO Journal* **10**, 1445-1457.

RICHARDS, W.L., TSUKADA, Y. & POTTER, V.R. (1982). Phenotypic diversity of gamma-glutamyltranspeptidase activity and protein secretion in hepatoma cell lines. *Cancer Research* **42**, 1374-1383.

RIJNTJES, P.J.M., MOSHAGE, H.J. & YAP, S.H. (1988). In vitro infection of primary cultures of cryopreserved adult human hepatocytes with hepatitis B virus. *Virus Research* **10**, 95-110.

ROBINSON, W.S., CLAYTON, D.A. & GREENMAN, R.L. (1974). DNA of a human hepatitis B virus candidate. *Journal of Virology* **14**, 384-391.

ROBINSON, W.S. (1977). The genome of hepatitis B virus. *Annual Review of Microbiology* **31**, 357-377.

ROBINSON, W.S. & GREENMAN, R.L. (1974). DNA polymerase in the core of the human hepatitis B virus candidate. *Journal of Virology* **13**, 1231-1236.

ROBINSON, W.S. & LUTWICK, L.I. (1976). The virus of hepatitis, Type B. *New England Journal of Medicine* 1168-1175.

ROMET-LEMONNE, J.L., ELFASSI, E., HASELTINE, W.A. & ESSEX, M. (1983a). Infection of bone marrow cells by hepatitis B virus. *Lancet* 2, 732.

ROMET-LEMONNE, J-L., FRANCES-MCLANE, M., ELFASSI, E., HASELTINE, W.A., AZOCAR, J. & ESSEX, M. (1983b). Hepatitis B virus infection in cultured human lymphoblastoid cells. *Science* 221, 667-669.

ROOSSINCK, M.J. & SIDDIQUI, A. (1987). In vivo phosphorylation and protein analysis of hepatitis B virus core antigen. *Journal of Virology* 61, 955-961.

ROSENFELD, M.G. (1991). POU-domain transcription factors: pou-er-ful developmental regulators. *Genes and Development* 5, 897-907.

ROY, A.L., MEISTERERNST, M., POGNONEC, P. & ROEDER, R.G. (1991). Cooperative interaction of an initiator-binding transcription initiation factor and the helix-loop-helix activator USF. *Nature, London* 354, 245-248.

RUVKUN, G. & FINNEY, M. (1991). Regulation of transcription and cell identity by POU domain proteins. *Cell* 64, 475-478.

RYFFEL, G.U., KUGLER, W., WAGNER, U. & KALING, M. (1989). Liver cell specific gene transcription in vitro: the promoter elements HP1 and TATA box are necessary and sufficient to generate a liver-specific promoter. *Nucleic Acids Research* 17, 939-953.

SAITO, I., OYA, Y. & SHIMOJO, H. (1986). Novel RNA family structure of hepatitis B virus expressed in human cells, using a helper-free adenovirus vector. *Journal of Virology* 58, 554-560.

SAMBROOK, J., FRITSCH, E.F. & MANIATIS, T. (1989). *Molecular Cloning: a Laboratory Manual*. New York: Cold Spring Harbor Laboratory Press.

SANGER, F., NICKLEN, S. & COULSON, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences, U.S.A.* 74, 5463-5467.

SATTLER, F. & ROBINSON, W.S. (1979). Hepatitis B viral DNA molecules have cohesive ends. *Journal of Virology* 32, 226-233.

SAWADAISHI, K., MORINAGA, T. & TAMAOKI, T. (1988). Interaction of a hepatoma-specific nuclear factor with transcription-regulatory sequences of the human α-fetoprotein and albumin genes. *Molecular and Cellular Biology* **8**, 5179-5187.

SAYRE, M.H., TSCHOCHNER, H. & KORNBERG, R.D. (1992a). Reconstitution of transcription with five purified initiation factors and RNA polymerase II from *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* **267**, 23376-23382.

SAYRE, M.H., TSCHOCHNER, H. & KORNBERG, R.D. (1992b). Purification and properties of *Saccharomyces cerevisiae* RNA polymerase II general initiation factor a. *Journal of Biological Chemistry* **267**, 23383-23387.

SCHLICHT, H-J., RADZIWILL, G. & SCHALLER, H. (1989). Synthesis and encapsidation of duck hepatitis B virus reverse transcriptase do not require formation of core-polymerase fusion proteins. *Cell* **56**, 85-92.

SCHOLER, H.R., RUPPERT, S., SUZUKI, N., CHOWDHURY, K. & GRUSS, P. (1990). New type POU domain in germ line-specific protein Oct-4. *Nature, London* **344**, 435-439.

SCHORPP, M., KUGLER, W., WAGNER, U. & RYFFEL, G.U. (1988). Hepatocyte-specific promoter element HP1 of the Xenopus albumin gene interacts with transcription factors of mammalian hepatocytes. *Journal of Molecular Biology* **202**, 307-320.

SCOTT, M.P., TAMKUN, J.W. & HARTZELL, G.W.,III (1993). The structure and function of the homeodomain. *Biochimica et Biophysica Acta* **989**, 25-48.

SEARLE, P.F., DAVIDSON, B.L., STUART, G.W., WILKIE, T.M., NORSTEDT, G. & PALMITER, R.D. (1984). Regulation, linkage, and sequence of mouse metallothionien I and II genes. *Molecular and Cellular Biology* **4**, 1221-1230.

SEEGER, C., GANEM, D. & VARMUS, H.E. (1986). Biochemical and genetic evidence for the hepatitis B virus replication strategy. *Science* **232**, 447-484.

SEEGER, C., BALDWIN, B. & TENNANT, B.C. (1989). Expression of infectious woodchuck hepatitis virus in murine and avian fibroblasts. *Journal of Virology* **63**, 4665-4669.

SEEGER, C. & MARAGOS, J. (1990). Identification and characterization of the woodchuck hepatitis virus origin of DNA replication. *Journal of Virology* **64**, 16-23.

SEIPEL, K., GEORGIEV, O. & SCHAFFNER, W. (1992). Different activation domains stimulate transcription from remote ('enhancer') and proximal ('promoter') positions. *EMBO Journal* 11, 4961-4968.

SELLS, M.A., CHEN, M-L. & ACS, G. (1987). Production of hepatitis B virus particles in HepG2 cells transfected with cloned hepatitis B virus DNA. *Proceedings of the National Academy of Sciences, U.S.A.* 84, 1005-1009.

SELLS, M.A., ZELENT, A.Z., SHVARTSMAN, M. & ACS, G. (1988). Replicative intermediates of hepatitis B virus in HepG2 cells that produce infectious virions. *Journal of Virology* 62, 2836-2844.

SETO, E., YEN, T.S.B., PETERLIN, B.M. & OU, J-H. (1988). Trans-activation of the human immunodeficiency virus long terminal repeat by the hepatitis B virus X protein. *Proceedings of the National Academy of Sciences, U.S.A.* 85, 8286-8290.

SETO, E., ZHOU, D-X., PETERLIN, B.M. & YEN, T.S.B. (1989). Trans-activation by the hepatitis B virus X protein shows cell-type specificity. *Virology* 173, 764-766.

SETO, E., MITCHELL, P.J. & YEN, T.S.B. (1990). Transactivation by the hepatitis B virus X protein depends on AP-2 and other transcription factors. *Nature, London* 344, 72-74.

SHAFRITZ, D.A., SHOUVAL, D., SHERMAN, H.I., HADZIYANNIS, S.J. & KEW, M.C. (1981). Integration of hepatitis B virus DNA into the genome of liver cells in chronic liver disease and hepatocellular carcinoma. *New England Journal of Medicine* 305, 1067-1073.

SHAUL, Y., RUTTER, W.J. & LAUB, O. (1985). A human hepatitis B viral enhancer element. *EMBO Journal* 4, 427-430.

SHAUL, Y., BEN LEVY, R. & DE MEDINA, T. (1986). High affinity binding site for nuclear factor I next to the hepatitis B virus S gene promoter. *EMBO Journal* 5, 1967-1971.

SHAUL, Y. & BEN LEVY, R. (1987). Multiple nuclear proteins in liver cells are bound to hepatitis B virus enhancer element and its upstream sequences. *EMBO Journal* 6, 1913-1920.

SHEPHERD, J.C.W., MCGINNIS, W., CARRASCO, A.E., DEROBERTIS, E.M. & GEHRING, W.J. (1984). Fly and frog homeo domains show homologies with yeast mating type regulatory proteins. *Nature, London* **310**, 70-71.

SHIH, C., LI, L-S., ROYCHOUDHURY, S. & HO, M-H. (1989). In vitro propagation of human hepatitis B virus in a rat hepatoma cell line. *Proceedings of the National Academy of Sciences, U.S.A.* **86**, 6323-6327.

SHIMODA, T., SHIKATA, T., KARASAWA, T., TSUKAGOSHI, S., YOSHIMURA, M. & SAKURAI, I. (1981). Light microscopic localization of hepatitis B virus antigens in the human pancreas. *Gastroenterology* **81**, 998-1005.

SIDDIQUI, A., JAMEEL, S. & MAPOLES, J. (1986). Transcriptional control elements of hepatitis B surface antigen gene. *Proceedings of the National Academy of Sciences, U.S.A.* **83**, 566-570.

SIDDIQUI, A,, JAMEEL, S. & MAPOLES, J. (1987). Expression of the hepatitis B virus X gene in mammalian cells. *Proceedings of the National Academy of Sciences, U.S.A.* **84**, 2513-2517.

SIDDIQUI, A., GAYNOR, R., SRINIVASAN, A., MAPOLES, J. & FARR, R.W. (1989). Trans-activation of viral enhancers including long terminal repeat of the human immunodeficiency virus by the hepatitis B virus X protein. *Virology* **169**, 479-484.

SIMON, K., LINGAPPA, V.R. & GANEM, D. (1988). Secreted hepatitis B surface antigen polypeptides are derived from a transmembrane precursor. *Journal of Cell Biology* **107**, 2163-2168.

SIMONSEN, C.C. & LEVINSON, A.D. (1983). Analysis of processing and polyadenylation signals of the hepatitis B virus surface antigen gene by using Simian virus 40-hepatitis B virus chimeric plasmids. *Molecular and Cellular Biology* **3**, 2250-2258.

SMALE, S.T., SCHMIDT, M.C., BERK, A.J. & BALTIMORE, D. (1990). Transcriptional activation by Sp1 as directed through TATA or initiator: Specific requirement for mammalian transcription factor IID. *Proceedings of the National Academy of Sciences, U.S.A.* **87**, 4509-4513.

SMALE, S.T. & BALTIMORE, D. (1989). The "initiator" as a transcription control element. *Cell* 57, 103-113.

SOMLO, F. (1989). Localization of hepatitis B virus in a primary testicular cancer. *Archives of Pathology and Laboratory Medicine* 113, 1184-1186.

SORGE, J., WRIGHT, D., ERDMAN, V.D. & CUTTING, A.E. (1984). Amphotropic retrovirus vector system for human cell gene transfer. *Molecular and Cellular Biology* 4, 1730-1737.

SPANDAU, D.F. & LEE, S-D. (1988). Trans-activation of viral enhancers by the hepatitis B virus X protein. *Journal of Virology* 62, 427-434.

SPRENGEL, R., KUHN, C., WILL, H. & SCHALLER, H. (1985). Comparative sequence analysis of duck and human hepatitis B virus genomes. *Journal of Medical Virology* 15, 323-333.

SPRENGEL, R., KALETA, E.F. & WILL, H. (1988). Isolation and characterization of a hepatitis B virus endemic in herons. *Journal of Virology* 62, 3832-3839.

STANDRING, D.N., RUTTER, W.J., VARMUS, H.E. & GANEM, D. (1984). Transcription of the hepatitis B surface antigen gene in cultured murine cells initiates within the presurface region. *Journal of Virology* 50, 563-571.

STANDRING, D.N., OU, J-H., MASIARZ, F.R. & RUTTER, W.J. (1988). A signal peptide encoded within the precore region of hepatitis B virus directs the secretion of heterogeneous population of e antigens in Xenopus oocytes. *Proceedings of the National Academy of Sciences, U.S.A.* 85, 8405-8409.

STEIN, O., FAINARU, M. & STEIN, Y. (1972). Visualization of virus-like particles in endoplasmic reticulum of hepatocytes of Australia antigen carriers. *Laboratory Investigation* 26, 262-269.

STEMLER, M., HESS, J., BRAUN, R., WILL, H. & SCHRODER, C.H. (1988). Serological evidence for expression of the polymerase gene of human hepatitis B virus in vivo. *Journal of General Virology* 69, 689-693.

STENLUND, A., LAMY, D., MORENO-LOPEZ, J., AHOLA, H., PETTERSON, U. & TIOLLAIS, P. (1983). Secretion of the hepatitis B virus surface antigen from mouse cells using an extra-chromosomal eucaryotic vector. *EMBO Journal* 2, 669-673.

STERN, S., TANAKA, M. & HERR, W. (1989). The Oct-1 homoeodomain directs formation of a multiprotein-DNA complex with the HSV transactivator VP16. *Nature, London* **341**, 624-630.

STERN, S. & HERR, W. (1991). The herpes simplex virus trans-activator VP16 recognizes the Oct-1 homeo domain: evidence for a homeo domain recognition subdomain. *Genes and Development* **5**, 2555-2566.

STIBBE, W. & GERLICH, W.H. (1982). Variable protein composition of hepatitis B surface antigen from different donors. *Virology* **123**, 436-442.

STIBBE, W. & GERLICH, W.H. (1983). Structural relationships between minor and major proteins of hepatitis B surface antigen. *Journal of Virology* **46**, 626-628.

STRINGER, K.F., INGLES, C.J. & GREENBLATT, J. (1990). Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID. *Nature, London* **345**, 783-786.

STURM, R.A., DAS, G. & HERR, W. (1988). The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a homeo box subdomain. *Genes and Development* **2**, 1582-1599.

STURM, R.A. & HERR, W. (1988). The POU domain is a bipartite DNA-binding structure. *Nature, London* **336**, 601-604.

SU, T-S., LAI, C-H., HUANG, J-L., LIN, L-H., YAUK, Y-K., CHANG, C., LO, S.J. & HAN, S-H. (1989a). Hepatitis B virus transcript produced by RNA splicing. *Journal of Virology* **63**, 4011-4018.

SU, T-S., LUI, W-Y., LIN, L-H., HAN, S-H. & P'ENG, F-K. (1989b). Analysis of hepatitis B virus transcripts in infected human livers. *Hepatology* **9**, 180-185.

SUMMERS, J., O'CONNELL, A. & MILLMAN, I. (1975). Genome of hepatitis B virus: Restriction enzyme cleavage and structure of DNA extracted from Dane particles. *Proceedings of the National Academy of Sciences, U.S.A.* **72**, 4597-4601.

SUMMERS, J., SMOLEC, J.M. & SNYDER, R. (1978). A virus similar to human hepatitis B virus associated with hepatitis and hepatoma in woodchucks. *Proceedings of the National Academy of Sciences, U.S.A.* **75**, 4533-4537.

SUMMERS, J. & MASON, W.S. (1982). Replication of the genome of a hepatitis B-like virus by reverse transcription of an RNA intermediate. *Cell* 29, 403-415.

SUREAU, C., ROMET-LEMONNE, J-L., MULLINS, J.I. & ESSEX, M. (1986). Production of hepatitis B virus by a differentiated human hepatoma cell line after transfection with cloned circular HBV DNA. *Cell* 47, 37-47.

SUREAU, C., EICHBERG, J.W., HUBBARD, G.B., ROMET-LEMONNE, J.L. & ESSEX, M. (1988). A molecularly cloned hepatitis B virus produced in vitro is infectious in a chimpanzee. *Journal of Virology* 62, 3064-3067.

SUTNICK, A.I., LONDON, W.T. & BLUMBERG, B.S. (1967). Australia antigen, Down's syndrome and hepatitis. *Journal of Clinical Investigation* 46, 1122.

SUTNICK, A.I., LONDON, W.T., GERSTLEY, B.J.S., CRONLUND, M.M. & BLUMBERG, B.S. (1968). Anicteric hepatitis associated with Australia antigen. *Journal of the American Medical Association* 205, 670-674.

SZMUNESS, W., MUCH, M.I., PRINCE, A.M., HOOFNAGLE, J.H., CHERUBIN, C.E., HARLEY, E.J. & BLOCK, G.H. (1975). On the role of sexual behavior in the spread of hepatitis B infection. *Annals of Internal Medicine* 83, 489-495.

TAGARIELLO, G., TRALDI, A., CAVALLETTO, D., PONTISSO, P. & ALBERTI, A. (1989). Hepatitis B Virus type 2 in Italian hemophiliacs. *Transfusion* 29, 558-559.

TAKADA, S., GOTOH, Y., HAYASHI, S., YOSHIDA, M. & KOIKE, K. (1990). Structural rearrangement of integrated hepatitis B virus DNA as well as cellular flanking DNA is present in chronically infected hepatic tissues. *Journal of Virology* 64, 822-828.

TAKAHASHI, K., MACHIDA, A., FUNATSU, G., NOMURA, M., USUDA, S., AOYAGI, S., TACHIBANA, K., MIYAMOTO, H., IMAI, M., NAKAMURA, T., MIYAKAWA, Y. & MAYUMI, M. (1983). Immunochemical structure of hepatitis B e antigen in the serum. *Journal of Immunology* 130, 2903-2907.

TANESE, N., PUGH, B.F. & TJIAN, R. (1991). Coactivators for a proline-rich activator purified from the multisubunit human TFIID complex. *Genes and Development* 5, 2212-2224.

THEILMANN, L., HOFMANN, W.J. & PFAFF, E. (1989a). Isolation and characterization of cytoplasmic and nuclear particles of hepatitis B virus. *Virus Research* **14**, 225-234.

THEILMANN, L., PFAFF, E., KOMMERELL, B., GMELIN, K., SCHALLER, H. & SALFELD, J. (1989b). Detection of hepatitis B virus core gene products in sera and liver of HBV-infected individuals. *Journal of Hepatology* **8**, 77-85.

THOMAS, H.C., JACYNA, M., WATERS, J. & MAIN, J. (1988). Virus-host interaction in chronic hepatitis B virus infection. *Seminars in Liver Disease* **8**, 342-349.

THUNG, S.N. & GERBER, M. (1981). Presence of receptors for polymerized albumin in HBs Ag-containing hepatocytes and hepatoma cell line. *Hepatology* **1**, 132-136.

TIAN, J.-M. & SCHIBLER, U. (1991). Tissue-specific expression of the gene encoding hepatocyte nuclear factor 1 may involve hepatocyte nuclear factor 4. *Genes and Development* **5**, 2225-2234.

TIOLLAIS, P., POURCEL, C. & DEJEAN, A. (1985). The hepatitis B virus. *Nature, London* **317**, 489-495.

TOGNONI, A., CATTANEO, R., SERFLING, E. & SCHAFFNER, W. (1985). A novel expression selection approach allows precise mapping of the hepatitis B virus enhancer. *Nucleic Acids Research* **13**, 7457-7472.

TOH, H., HAYASHIDA, H. & MIYATA, T. (1983). Sequence homology between retroviral reverse transcriptase and putative polymerases of hepatitis B virus and cauliflower mosaic virus. *Nature, London* **305**, 827-829.

TONG, M.J., THURSBY, M., RAKELA, J., MCPEAK, C., EDWARDS, V.M. & MOSLEY, J.W. (1981). Studies on the maternal-infant transmission of the viruses which cause acute hepatitis. *Gastroenterology* **80**, 999-1004.

TONIATTI, C., MONACI, P., NICOSIA, A., CORTESE, R. & CILIBERTO, G. (1993). A bipartite activation domain is responsible for the activity of transcription factor HNF1/LFB1 in cells of hepatic and nonhepatic origin. *DNA and Cell Biology* **12**, 199-208.

TREININ, M. & LAUB, O. (1987). Identification of a promoter element located upstream from the hepatitis B virus X gene. *Molecular and Cellular Biology* **7**, 545-548.

TRIEZENBERG, S.J., KINGSBURY, R.C. & MCKNIGHT, S.L. (1988a). Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression. *Genes and Development* **2**, 718-729.

TRIEZENBERG, S.J., LAMARCO, K.L. & MCKNIGHT, S.L. (1988b). Evidence of DNA:protein interactions that mediate HSV-1 immediate early gene activation by VP16. *Genes and Development* **2**, 730-742.

TRONCHE, F. & YANIV, M. (1992). HNF1, a homeoprotein member of the hepatic transcription regulatory network. *Bioessays* **14**, 579-587.

TRUJILLO, M.A., LETOVSKY, J., MAGUIRE, H.F., LOPEZ-CABRERA, M. & SIDDIQUI, A. (1991). Functional analysis of a liver-specific enhancer of the hepatitis B virus. *Proceedings of the National Academy of Sciences, U.S.A.* **88**, 3797-3801.

TSCHOCHNER, H., SAYRE, M.H., FLANAGAN, P.M., FEAVER, W.J. & KORNBERG, R.D. (1992). Yeast RNA polymerase II initiation factor e: Isolation and identification as the functional counterpart of human transcription factor IIB. *Proceedings of the National Academy of Sciences, U.S.A.* **89**, 11292-11296.

TSURIMOTO, T., FUJIYAMA, A. & MATSUBARA, K. (1987). Stable expression and replication of hepatitis B virus genome in an integrated state in a human hepatoma cell line transfected with the cloned viral DNA. *Proceedings of the National Academy of Sciences, U.S.A.* **84**, 444-448.

TSUTSUMI, K.I., ITO, K. & ISHIKAWA, K. (1989). Developmental appearance of transcription factors that regulate liver-specific expression of the aldolase B gene. *Molecular and Cellular Biology* **9**, 4923-4931.

TUR KASPA, R., BURK, R.D., SHAUL, Y. & SHAFRITZ, D.A. (1986). Hepatitis B virus DNA contains a glucocorticoid-responsive element. *Proceedings of the National Academy of Sciences, U.S.A.* **83**, 1627-1631.

TUR KASPA, R., SHAUL, Y., MOORE, D.D., BURK, R.D., OKRET, S., POELLINGER, L. & SHAFRITZ, D.A. (1988). The glucocorticoid receptor recognizes a specific nucleotide sequence in hepatitis B virus DNA causing increased activity of the HBV enhancer. *Virology* **167**, 630-633.

TURNER, R. & TJIAN, R. (1989). Leucine repeats and an adjacent DNA binding domain mediate the formation of functional cFos-cJun heterodimers. *Science* 243, 1689-1694.

TUTTLEMAN, J.S., POURCEL, C. & SUMMERS, J. (1986). Formation of the pool of covalently closed circular viral DNA in hepadnavirus-infected cells. *Cell* 47, 451-460.

TWU, J-S. & ROBINSON, W.S. (1989). Hepatitis B virus X gene can transactivate heterologous viral sequences. *Proceedings of the National Academy of Sciences, U.S.A.* 86, 2046-2050.

TWU, J-S. & SCHLOEMER, R.H. (1987). Transcriptional trans-activating function of hepatitis B virus. *Journal of Virology* 61, 3448-3453.

TWU, J.S., CHU, K. & ROBINSON, W.S. (1989a). Hepatitis B virus X gene activates kappa B-like enhancer sequences in the long terminal repeat of human immunodeficiency virus 1. *Proceedings of the National Academy of Sciences, U.S.A.* 86, 5168-5172.

TWU, J.S., ROSEN, C.A., HASELTINE, W.A. & ROBINSON, W.S. (1989b). Identification of a region within the human immunodeficiency virus type 1 long terminal repeat that is essential for transactivation by the hepatitis B virus gene X. *Journal of Virology* 63, 2857-2860.

UEDA, K., TSURIMOTO, T. & MATSUBARA, K. (1991). Three envelope proteins of hepatitis B virus: Large S, middle S, and major S proteins needed for the formation of Dane particles. *Journal of Virology* 65, 3521-3529.

VACHER, J. & TILGHMAN, S.M. (1990). Dominant negative regulation of the mouse α-fetoprotein gene in adult liver. *Science* 250, 1732-1735.

VALENZUELA, P., GRAY, P., QUIROGA, M., ZALDIVAR, J., GOODMAN, H.M. & RUTTER, W.J. (1979). Nucleotide sequence of the gene coding for the major protein of hepatitis B virus surface antigen. *Nature, London* 280, 815-819.

VALENZUELA, P., QUIROGA, M., ZALDIVAR, J., GRAY, P. & RUTTER, W.J. (1980). The nucleotide sequence of the hepatitis B viral genome and the identification of the major viral genes. In *Animal Virus Genetics*, pp. 57-70. Edited by B. Fields, R. Jaenisch & C. F. Fox. New York: Academic Press, Inc..

VAN ZONNEVELD, A.J., CURRIDEN, S.A. & LOSKUTOFF, D.J. (1988). Type 1 plasminogen activator inhibitor gene: Functional analysis and glucocorticoid regulation of its promoter. *Proceedings of the National Academy of Sciences, U.S.A.* **85**, 5525-5529.

VANNICE, J.L. & LEVINSON, A.D. (1988). Properties of the human hepatitis B virus enhancer: position effects and cell-type nonspecificity. *Journal of Virology* **62**, 1305-1313.

VARRIJZER, C.P., ALKEMA, M.J., VAN WEPEREN, W.W., VAN LEEUWEN, H.C., STRATING, M.J.J. & VAN DER VLIET, P.C. (1992). The DNA binding specificity of the bipartite POU domain and its subdomains. *EMBO Journal* **11**, 4993-5003.

VAUDIN, M., WOLSTENHOLME, A.J., TSIQUAY, K.N., ZUCKERMAN, A.J. & HARRISON, T.J. (1988). The complete nucleotide sequence of the genome of a hepatitis B virus isolated from a naturally infected chimpanzee. *Journal of General Virology* **69**, 1383-1389.

VAULONT, S., PUZENAT, N., KAHN, A. & RAYMONDJEAN, M. (1989). Analysis by cell-free transcription of the liver-specific pyruvate kinase gene promoter. *Molecular and Cellular Biology* **9**, 4409-4415.

VERRIJZER, C.P., KAL, A.J. & VAN DER VLIET, P.C. (1990a). The oct-1 homeo domain contacts only part of the octamer sequence and full oct-1 DNA-binding activity requires the POU-specific domain. *Genes and Development* **4**, 1964-1974.

VERRIJZER, C.P., KAL, A.J. & VAN DER VLIET, P.C. (1990b). The DNA binding domain (POU domain) of transcription factor oct-1 suffices for stimulation of DNA replication. *EMBO Journal* **9**, 1883-1888.

VERRIJZER, C.P., VAN OOSTERHOUT, J.A.W.M., VAN WEPEREN, W.W. & VAN DER VLIET, P.C. (1991). POU proteins bend DNA via the POU-specific domain. *EMBO Journal* **10**, 3007-3014.

VERRIJZER, C.P., VAN OOSTERHOUT, J.A.W.M. & VAN DER VLIET, P.C. (1992). The Oct-1 POU domain mediates interactions between Oct-1 and other POU proteins. *Molecular and Cellular Biology* **12**, 542-551.

VIEIRA, J. & MESSING, J. (1982). The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**, 259-268.

WAISMAN, A., ALONI, Y. & LAUB, O. (1990). *In vitro* regulation of human hepatitis B virus core gene transcription. *Virology* **177**, 737-744.

WANDS, J.R., FUJITA, Y.K., ISSELBACHER, K.J., DEGOTT, C., SCHELLEKENS, H., DAZZA, M-CH., THIERS, V., TIOLLAIS, P. & BRECHOT, C. (1986). Identification and transmission of hepatitis B virus-related variants. *Proceedings of the National Academy of Sciences, U.S.A.* **83**, 6608-6612.

WANG, Y., CHEN, P., WU, X., SUN, A.-L., WANG, H., ZHU, Y.-A. & LI, Z.-P. (1990). A new enhancer element, ENII, identified in the X gene of hepatitis B virus. *Journal of Virology* **64**, 3977-3981.

WEIMER, T., SALFELD, J. & WILL, H. (1987). Expression of the hepatitis B virus core gene in vitro and in vivo. *Journal of Virology* **61**, 3109-3113.

WEIMER, T., WEIMER, K., TU, Z-X., JUNG, M-C., PAPE, G.R. & WILL, H. (1989). Immunogenicity of human hepatitis B virus P-gene derived proteins. *Journal of Immunology* **143**, 3750-3756.

WEINZIERL, R.O.J., DYNLACHT, B.D. & TJIAN, R. (1993). Largest subunit of Drosophila transcription factor IID directs assembly of a complex containing TBP and a coactivator. *Nature, London* **362**, 511-517.

WEIS, L. & REINBERG, D. (1992). Transcription by RNA polymerase II: Initiator-directed formation of transcription-competent complexes. *FASEB Journal* **6**, 3300-3309.

WEISIGER, R., GOLLAN, J.L. & OCKNER, R. (1981). Receptor for albumin on the liver cell surface may mediate uptake of fatty acids and other albumin-bound substances. *Science* **211**, 1048-1051.

WHITE, R.A.H. & WILCOX, M. (1984). Protein products of the bithorax complex in Drosophila. *Cell* **39**, 163-171.

WILL, H., CATTANEO, R., KOCH, H-G., DARAI, G. & SCHALLER, H. (1982). Cloned HBV DNA causes hepatitis in chimpanzees. *Nature, London* **299**, 740-742.

WILL, H., CATTANEO, R., DARAI, G., DEINHARDT, F., SCHELLEKENS, H. & SCHALLER, H. (1985). Infectious hepatitis B virus from cloned DNA of known nucleotide sequence. *Proceedings of the National Academy of Sciences, U.S.A.* **82**, 891-895.

WILL, H., REISER, W., WEIMER, T., PFAFF, E., BUSCHER, M., SPRENGLE, R., CATTANEO, R. & SCHALLER, H. (1987). Replication strategy of human hepatitis B virus. *Journal of Virology* **61**, 904-911.

WILLIAMS, T., ADMON, A., LUSCHER, B. & TJIAN, R. (1988). Cloning and expression of AP-2, a cell-type-specific transcription factor that activates inducible enhancer elements. *Genes and Development* **2**, 1557-1569.

WILLIAMS, T. & TJIAN, R. (1991). Analysis of the DNA-binding and activation properties of the human transcription factor AP-2. *Genes and Development* **5**, 670-682.

WOLBERGER, C., VERSHON, A.K., LIU, B., JOHNSON, A.D. & PABO, C.O. (1991). Crystal structure of a MATα2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. *Cell* **67**, 517-528.

WU, J-S., KO, Y-C. & LIU, W.T. (1988). Hepatitis B virus type 2. *Lancet* **1**, 990.

WUARIN, J., MUELLER, C. & SCHIBLER, U. (1990). A ubiquitous CCAAT factor is required for efficient *in vitro* transcription from the mouse albumin promoter. *Journal of Molecular Biology* **214**, 865-874.

YAGINUMA, K., SHIRAKATA, Y., KOBAYASHI, M. & KOIKE, K. (1987). Hepatitis B virus (HBV) particles are produced in a cell culture system by transient expression of transfected HBV DNA. *Proceedings of the National Academy of Sciences, U.S.A.* **84**, 2678-2682.

YAGINUMA, K. & KOIKE, K. (1989). Identification of a promoter region for 3.6-kilobase mRNA of hepatitis B virus and specific cellular binding protein. *Journal of Virology* **63**, 2914-2920.

YAMADA, G., SAKAMOTO, Y., MIZUNO, M., NISHIHARA, T., KOBAYASHI, T., TAKAHASHI, T. & NAGASHIMA, H. (1982). Electron and immunoelectron microscopic study of Dane particle formation in chronic hepatitis B virus infection. *Gastroenterology* **83**, 348-356.

YAMADA, G. & NAKANE, P.K. (1977). Hepatitis B core and surface antigens in liver tissue. *Laboratory Investigation* **36**, 649-659.

YAMAKI, M., OHORI, H., ONODERA, S., ISHIDA, N. & MAEDA, H. (1982). Circular dichroism and biochemical properties of the hepatitis B virus core antigen. *Biochimica et Biophysica Acta* 706, 165-173.

YEE, J-K. (1989). A liver-specific enhancer in the core promoter region of human hepatitis B virus. *Science* 246, 658-661.

YOKOSUKA, O., OMATA, M., IMAZEKI, F., ITO, Y. & OKUDA, K. (1986). Hepatitis B virus RNA transcripts and DNA in chronic liver disease. *New England Journal of Medicine* 315, 1187-1192.

YONEYAMA, T., AKATSUKA, T. & MIYAMURA, T. (1988). Stable expression of the hepatitis B virus surface antigen containing Pre-S2 protein in mouse cells using a bovine papillomavirus vector. *Journal of General Virology* 69, 1931-1939.

YUH, C.-H., CHANG, Y.-L. & TING, L.-P. (1992). Transcriptional regulation of precore and pregenomic RNAs of hepatitis B virus. *Journal of Virology* 66, 4073-4084.

YUH, C.-H. & TING, L.-P. (1990). The genome of hepatitis B virus contains a second enhancer: Cooperation of two elements within this enhancer is required for its function. *Journal of Virology* 64, 4281-4287.

YUH, C.-H. & TING, L.-P. (1991). C/EBP-like proteins binding to the functional box-α and box-β of the second enhancer of hepatitis B virus. *Molecular and Cellular Biology* 11, 5044-5052.

YUH, C.-H. & TING, L.-P. (1993). Differentiated liver cell specificity of the second enhancer of hepatitis B virus. *Journal of Virology* 67, 142-149.

ZAHM, P., HOFSCHNEIDER, P.H. & KOSHY, R. (1988). The HBV X-ORF encodes a transactivator: A potential factor in viral hepatocarcinogenesis. *Oncogene* 3, 169-177.

ZAWEL, L. & REINBERG, D. (1993). Initiation of transcription by RNA polymerase II: a multi-step process. In *Progress in nucleic acid research and molecular biology, vol. 44*, pp. 67-108. Edited by W. E. Cohn & K. Moldave. New York: Academic Press, Inc..

ZELENT, A.Z., SELLS, M.A., PRICE, P.M., MOHAMAD, A., ACS, G. & CHRISTMAN, J.K. (1987a). Murine cells carrying integrated tandem genomes of hepatitis B virus DNA transcribe RNAs from endogenous promoters on both viral

strands and express middle and major viral envelope proteins. *Journal of Virology* **61**, 1108-1115.

ZELENT, A.Z., SELLS, M.A., SHVARTSMAN, M., PRICE, P.M. & ACS, G. (1987b). Replicative intermediate of hepatitis B virus in transfected murine fibroblasts. *Journal of Virology* **61**, 2921-2923.

ZHANG, P., RANEY, A.K. & MCLACHLAN, A. (1992). Characterization of the hepatitis B virus X- and nucleocapsid gene transcriptional regulatory elements. *Virology* **191**, 31-41.

ZHANG, P., RANEY, A.K. & MCLACHLAN, A. (1993). Characterization of functional Sp1 transcription factor binding sites in the hepatitis B virus nucleocapsid promoter. *Journal of Virology* **67**, 1472-1481.

ZHOU, D.-X. & YEN, T.S.B. (1990). Differential regulation of the hepatitis B virus surface gene promoters by a second viral enhancer. *Journal of Biological Chemistry* **265**, 20731-20734.

ZHOU, D.-X. & YEN, T.S.B. (1991). The ubiquitous transcription factor Oct-1 and the liver-specific factor HNF-1 are both required to activate transcription of a hepatitis B virus promoter. *Molecular and Cellular Biology* **11**, 1353-1359.

ZOULIM, F., VITVITSKI, L., PICHOUD, C., ROUGIER, P., BOUFFARD, P. & TREPO, C. (1988). Expression of pre-S proteins by peripheral blood mononuclear cells (PBMC) in chronic hepatitis: A relevant non invasive approach to HBV replication. *Hepatology* **8**, 1389.